Cooperative Algorithms for MIMO Amplify-and-Forward Relay Networks

Kien T. Truong, *Student Member, IEEE,* Philippe Sartori, and

Robert W. Heath, Jr.*, *Fellow, IEEE*

Abstract

Interference alignment is a signaling technique that provides high multiplexing gain in the interference channel. It can be extended to multi-hop interference channels, where relays aid transmission between sources and destinations. In addition to coverage extension and capacity enhancement, relays increase the multiplexing gain in the interference channel. In this paper, three cooperative algorithms are proposed for a multiple-antenna amplify-and-forward (AF) relay interference channel. The algorithms design the transmitters and relays so that interference at the receivers can be aligned and canceled. The first algorithm minimizes the sum power of enhanced noise from the relays and interference at the receivers. The second and third algorithms rely on a connection between mean square error and mutual information to solve the end-to-end sum-rate maximization problem with either equality or inequality power constraints via matrix-weighted sum mean square error minimization. Since we can find a globally optimal solution in each iteration, the resulting iterative algorithms are convergent. Simulations show that the proposed algorithms achieve higher end-to-end sum-rates and multiplexing gains that existing strategies for AF relays, decode-and-forward relays, and direct transmission. The first algorithm outperforms the other algorithms at high signal-to-noise ratio (SNR) but performs worse than them at low SNR. Thanks to power control, the third algorithm outperforms the second algorithm at the cost of additional overhead.

Index Terms

Interference alignment, relay-aided interference alignment, two-hop interference channel, relay interference channel, relay beamforming, joint source-relay design.

I. INTRODUCTION

Relay interference channels model networks where a stage of intermediate nodes, called relays, help multiple transmitters communicate with their receivers using shared radio resources [1]–[4]. Upcoming cellular standards are considering relay communication for coverage extension and capacity enhancement [5], [6]. Prior work, however, shows that single-antenna relays do not work well in the presence of co-channel interference [7], [8]. In this paper, we consider multiple-antenna relay systems to take the advantage of the interference management capability of multiple-input multiple-output (MIMO) communication. Many interference management strategies have been proposed for the MIMO single-hop interference channel [9]–[13]. Although these single-hop results can be applied separately for the transmitter-relay hop and for the relay-receiver hop, even higher sum-rates can be achieved if the relays are configured jointly [14], [15]. Obtaining the most from relay interference channels requires advanced interference management strategies that jointly configure the transmitters, relays, and receivers.

A general challenge to designing algorithms for the interference channel is that the sum capacity is unknown. The multiplexing gain of a network is a first-order approximation of its sum-capacity at high signal-to-noise ratio (SNR) [16]. Interference alignment is a multiplexing gain maximizing signaling technique for the single-hop interference channel [13], achieving the maximum number of degrees of freedom. The idea is to arrange the transmitted signals such that interference is constrained within only a portion of the signal space observed by each receiver, leaving the remaining portion for interference-free detection of the desired signal [9]. The maximum multiplexing gains achievable through interference alignment, however, depend on the characteristics of the interference channels. For a symmetric MIMO interference channel with constant channel coefficients, the maximum multiplexing gain is upper-bounded by the total number of antennas at a transmitter-receiver pair regardless the number of pairs [11], [12]. Note that the bound is tight in certain cases and corresponds to the total available spatial dimensions of a pair. Increasing the number of spatial dimensions in the network, using for example relays, is one way to improve the maximum achievable multiplexing gain.

Relays can be classified based on their signal processing operation, among which the most popular are decode-and-forward (DF - the relays decode the received signals then re-encode before retransmitting) and amplify-and-forward (AF - the relays apply linear signal processing to the signal before forwarding). Without decoding the received signals, AF relays need no knowledge of the codebooks used by the transmitters and likely have lower baseband complexity and fast signal processing. In addition, transparent

to the modulation and coding of the signals, AF relays are more suitable for applications in heterogeneous networks comprising many nodes of different complexity or even standards [17]. In this paper, we focus on a half-duplex MIMO AF relay interference channel. Since half-duplex relays cannot transmit and receive at the same time, they are more practical than full-duplex relays.

Several interference management strategies designed specifically for the one-way AF relay interference channel have been proposed [15], [18]–[32]. Although relays cannot improve the multiplexing gains of the single-antenna fully-connected interference channel with time-varying or frequency-selective channel coefficients [18], they are beneficial for reducing the number of independent channel extensions needed to align interference at the receivers [19]. Prior work often considers networks operating in special circumstances. It is assumed in [15], [20]–[24] that there are enough antennas at the relays to cancel all interference on the reception and then to nullify all interference on the retransmission, allowing multiplexing gains to scale linearly with the number of users. Other prior work considers only small networks with up to three pairs to derive some kind of closed-form strategies [25]–[27]. Prior work in [28] considers design problems with different objective functions including sum power minimization and minimum SINR maximization. In addition, the algorithms in [28] are applicable only for single-antenna receivers. Prior work in [30]–[32] develops noncooperative resource allocation strategies for AF relay networks while our work focuses on centralized algorithms for cooperative resource allocation. In this paper, we consider a general setting in the sense that we assume no special constraints on the number of wireless nodes or the number of antennas at a node. The closest AF relay model to ours is considered in [29], which is only for single-antenna transmitters and receivers. Further, they assume no crosslinks from relays to receivers, resulting in an oversimplified design problem.

Sum-rate maximization problems are nonconvex and NP-hard, i.e., their global optima cannot be found in a polynomial time. It is even challenging to find their good local optima corresponding to interference aligned solutions using gradient-based algorithms from arbitrary initializations because those solutions have very narrow regions of attraction [33]. Thus, we propose to formulate three new design problems that have exactly the same constraints with sum-rate maximization problems but with better-behaved objective functions to find high-quality solutions in terms of sum-rate maximization. Based on the observation that interference alignment solutions make total leakage power go to zero [19], [34], [35], we formulate a problem that aims at minimizing the sum power of interference and enhanced noise from the relays. Based on a relationship between achievable rates and mean squared error (MSE), we formulate two matrix-

weighted sum-MSE minimization problems, either without power control or with power control. The key is that they have the same stationary points as their corresponding end-to-end sum-rate maximization problems. Although the newly formulated optimization problems are still nonconvex and NP-hard, they may be easier to solve. Next, we propose to adopt an alternating minimization approach [36], [37] to develop iterative algorithms for solving the newly formulated problems. In each iteration, all but one variable is fixed and we focus on designing the remaining variable by solving a single-variable optimization problem obtained from the corresponding original multi-variable problems. Since we are able to find a global optimum for the single-variable optimization problem in each iteration, the proposed algorithms are guaranteed to converge. Note that the power constraints at the relays depend on both the transmit precoders at the transmitters and the processing matrices at the relays, adding more constraints to the design problems. Thus, it is not straightforward to extend the methods used for the single-hop design problems to solve the two-hop design problems. Our initial results in this paper were reported in [38]. Compared with [38], this paper presents three different algorithms, has more discussion of convergence and provides simulations that emphasize the achievable end-to-end sum-rates and multiplexing gains.

We use Monte Carlo simulation to evaluate the average end-to-end sum-rates and multiplexing gains achievable through the proposed algorithms. First, the numerical results confirm the convergence of the proposed algorithms as expected. Second, over the iterations of the total leakage minimization algorithm, the true interference dominates at the beginning but is canceled quickly; after that, the enhanced noise from the relays becomes dominant. This means that relay-aided interference alignment should take into account the enhanced noise from the relays. Third, the total leakage minimization algorithm achieves lower end-to-end sum-rates than the others at low-to-medium SNR values because it ignores the desired signal power and noise power at the receivers. Nevertheless, the MSE-based algorithms result in unfairness, i.e., some users have much smaller rates than the others. Thus, the MSE-based algorithms achieve lower end-to-end sum-rates and multiplexing gains than the total leakage minimization algorithm at high SNR. One reason for this is that the MSE-based algorithms may either turn off some data streams or nullify the desired signals to some receivers. Fourth, for fixed numbers of antennas at the transmitters and at the receivers, even with half-duplex loss, AF relays can provide larger end-to-end multiplexing gains than DF relays or direct transmissions. Finally, the results show that AF relays provide larger average achievable end-to-end sum-rates than do DF relays. The proposed algorithms also provide higher achievable end-to-end sum-rates than the existing AF relaying strategies that do not align interference at the receivers.

The organization of the remainder of this paper is as follows. Section II describes the system model. Section III formulates the end-to-end sum-rate maximization problems and presents our proposed approach. Section IV develops three cooperative algorithms that aim at finding high-quality solutions of the sum-rate maximization problems. Section V evaluates numerically the proposed algorithms. Section VI concludes this paper and suggests future research.

Notation: We use normal letters (e.g., a) for scalars, lowercase and uppercase boldface letters (e.g., $\mathbf{h}$ and $\mathbf{H}$) for column vectors and matrices. $\mathbf{I}_N$ and $\underline{0}_N$ are the identity matrix and all-zero matrices of size $N \times N$. $\nu_{\min}^n(\mathbf{A})$ gives the eigenvectors corresponding to the n smallest eigenvalues of $\mathbf{A}$. For a matrix $\mathbf{A}$, $\mathbf{A}^T$ is the transpose matrix, $\|\mathbf{A}\|_F^2$ the Frobenious norm, $\mathbf{A}^*$ the conjugate transpose, and $\mathrm{tr}(\mathbf{A})$ the trace. $\mathrm{vec}(\mathbf{A})$ denotes the vec operator to transform $\mathbf{A}$ into $\mathbf{a}$ while $\mathrm{vec}^{-1}(\mathbf{a})$ denotes the inverse operator. $\otimes$ is the Kronecker product. $\mathbb{E}[\cdot]$ is the statistical expectation operator. $()^{(n)}$ denotes iteration index. $()_\mathrm{T}$ is used for transmitters' parameters, $()_\mathrm{R}$ for receivers', and $()_\mathrm{X}$ for relays'.

II. SYSTEM MODEL

Consider a relay interference channel where M half-duplex AF relays aid the one-way communication between K pairs of transmitters and receivers, as illustrated in Fig. 1. Each transmitter has data for only one receiver and each receiver is served by only one transmitter. Each pair is assigned a unique index $k \in \mathcal{K} \triangleq \{1, \cdots, K\}$. Transmitter k has $N_{\mathrm{T},k}$ antennas while receiver k has $N_{\mathrm{R},k}$ antennas for $k \in \mathcal{K}$. Similarly, each relay is assigned a unique index $m \in \mathcal{M} \triangleq \{1, \cdots, M\}$. Relay m has $N_{\mathrm{X},m}$ antennas for $m \in \mathcal{M}$. The half-duplex relays cannot transmit and receive at the same time, thus the transmission procedure consists of two stages. In the first stage, the transmitters send data to the relays. In the second stage, the relays apply linear processing to the received signals and forward to the receivers. We assume the direct channels between the transmitters and the receivers are ignored by the second-stage receivers.
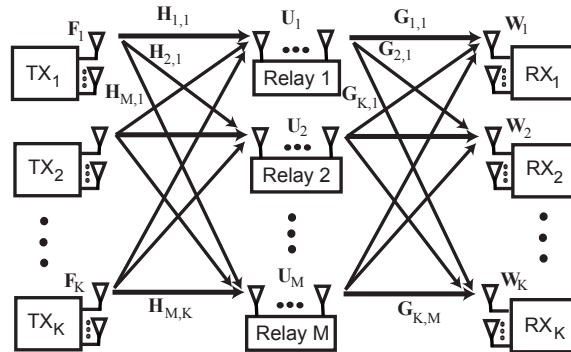

Fig. 1. A relay interference channel where M half-duplex AF relays aid the communication of K transmitter-receiver pairs.

We denote $\mathbf{H}_{m,k} \in \mathbb{C}^{N_{\mathrm{X},m} \times N_{\mathrm{T},k}}$ as the matrix channel from transmitter k to relay m and $\mathbf{G}_{k,m} \in \mathbb{C}^{N_{\mathrm{R},k} \times N_{\mathrm{X},m}}$ as the matrix channel from relay m to receiver k for $k \in \mathcal{K}$ and $m \in \mathcal{M}$. We assume that perfect and instantaneous knowledge of $\mathbf{H}_{m,k}$ and $\mathbf{G}_{k,m}$ for $k \in \mathcal{K}$ and $m \in \mathcal{M}$ is available at a central processing unit. Although this is a strict requirement, our results are still valuable since they show the substantial gains that can be achieved through coordination. Our results can be used as a benchmark for future work that makes more practical CSI assumptions.

Let $\mathbf{s}_k \in \mathbb{C}^{d_k \times 1}$ be the transmit symbol vector at transmitter k, where $d_k \le \min\{N_{\mathrm{T},k}, N_{\mathrm{R},k}\}$ is the number of data streams from transmitter k to receiver k for $k \in \mathcal{K}$. The transmit symbols are independent identically distributed (i.i.d.) such that $\mathbb{E}(\mathbf{s}_k\mathbf{s}_k^*) = \mathbf{I}_{d_k}$. Transmitter k uses a linear transmit precoder $\mathbf{F}_k \in \mathbb{C}^{N_{\mathrm{T},k} \times d_k}$ to map $\mathbf{s}_k$ to its transmit antennas. Let $p_{\mathrm{T},k}^{\max}$ be the maximum transmit power. The actual transmit power at transmitter k is $p_{\mathrm{T},k} = \mathrm{tr}(\mathbf{F}_k^*\mathbf{F}_k)$. Let $\mathbf{n}_{\mathrm{X},m}$ be spatially white, additive Gaussian noise at relay m with covariance $\mathbb{E}(\mathbf{n}_{\mathrm{X},m}\mathbf{n}_{\mathrm{X},m}^*) = \sigma_{\mathrm{X},m}^2\mathbf{I}_{N_{\mathrm{X},m}}$ for $m \in \mathcal{M}$. With perfect synchronization, relay m observes the following signal

$$\mathbf{y}_{\mathrm{X},m} = \sum_{k=1}^{K} \underbrace{\mathbf{H}_{m,k}\mathbf{F}_k}_{\boldsymbol{\mathcal{H}}_{m,k}} \mathbf{s}_k + \mathbf{n}_{\mathrm{X},m}. \tag{1}$$

Let $\mathbf{U}_m \in \mathbb{C}^{N_{\mathrm{X},m} \times N_{\mathrm{X},m}}$ be the processing matrix at relay m. The transmit signal at relay m is given by

$$\mathbf{x}_{\mathrm{X},m} = \mathbf{U}_m\mathbf{y}_{\mathrm{X},m} = \sum_{k=1}^{K} \mathbf{U}_m\boldsymbol{\mathcal{H}}_{m,k}\mathbf{s}_k + \mathbf{U}_m\mathbf{n}_{\mathrm{X},m}. \tag{2}$$

Relay m actually uses the following transmit power

$$p_{\mathrm{X},m} = \sum_{k=1}^{K} \mathrm{tr}(\mathbf{U}_m\boldsymbol{\mathcal{H}}_{m,k}\boldsymbol{\mathcal{H}}_{m,k}^*\mathbf{U}_m^*) + \sigma_{\mathrm{X},m}^2\,\mathrm{tr}(\mathbf{U}_m\mathbf{U}_m^*). \tag{3}$$

There are two possible types of power constraints at the relays: i) a set of individual power constraints at the relays and ii) a sum power constraint at all the relays. Individual relay power constraints are often considered in the cellular system literature [39], [40]. While a sum power constraint is often considered in the ad hoc network literature to extend the lifetime of battery-powered relays [41], [42]. Let $p_{\mathrm{X},m}^{\max}$ be the maximum transmit power at relay m and $p_{\mathrm{X}}^{\max}$ be the maximum sum transmit power at all the relays. When power control is considered, the individual relay power constraints are $p_{\mathrm{X},m} \le p_{\mathrm{X},m}^{\max}, \forall m \in \mathcal{M}$; whereas the sum relay power constraint is $\sum_{m=1}^{M} p_{\mathrm{X},m} \le p_{\mathrm{X}}^{\max}$. Without power control, the inequalities in the power constraint expressions are replaced by equalities. The following sections focus on the sum power constraint at the relays. Section IV-D discusses the applicability of individual relay power constraints while Section V simulates the impact of individual relay power constraints on achievable

end-to-end sum-rates.

Let $\mathbf{n}_{\mathrm{R},k}$ be spatially white, additive Gaussian noise at receiver k with covariance $\mathbb{E}(\mathbf{n}_{\mathrm{R},k}\mathbf{n}_{\mathrm{R},k}^*) = \sigma_{\mathrm{R},k}^2 \mathbf{I}_{N_{\mathrm{R},k}}$. We denote $\boldsymbol{\mathcal{G}}_{k,m} = \mathbf{G}_{k,m}\mathbf{U}_m$. Receiver k observes the following signal

$$\mathbf{y}_k = \sum_{m=1}^{M} \mathbf{G}_{k,m}\mathbf{x}_{\mathrm{X},m} + \mathbf{n}_{\mathrm{R},k} \tag{4}$$

$$= \sum_{q=1}^{K} \underbrace{\sum_{m=1}^{M} \boldsymbol{\mathcal{G}}_{k,m}\boldsymbol{\mathcal{H}}_{m,q}}_{\boldsymbol{\mathcal{T}}_{k,q}} \mathbf{s}_q + \sum_{m=1}^{M} \boldsymbol{\mathcal{G}}_{k,m}\mathbf{n}_{\mathrm{X},m} + \mathbf{n}_{\mathrm{R},k}, \tag{5}$$

where $\boldsymbol{\mathcal{T}}_{k,q}$ is the effective end-to-end channel from transmitter q to receiver k for $k, q \in \mathcal{K}$. Applying a linear receive filter $\mathbf{W}_k \in \mathbb{C}^{N_{\mathrm{R},k} \times d_k}$ to $\mathbf{y}_k$, receiver k obtains

$$\bar{\mathbf{y}}_k = \underbrace{\mathbf{W}_k^* \boldsymbol{\mathcal{T}}_{k,k}\mathbf{s}_k}_{\text{desired signal}} + \underbrace{\sum_{\substack{q=1 \\ q \neq k}}^{K} \mathbf{W}_k^* \boldsymbol{\mathcal{T}}_{k,q}\mathbf{s}_q}_{\text{interference}} + \underbrace{\sum_{m=1}^{M} \mathbf{W}_k^* \boldsymbol{\mathcal{G}}_{k,m}\mathbf{n}_{\mathrm{X},m}}_{\text{enhanced noise from relays}} + \underbrace{\mathbf{W}_k^*\mathbf{n}_{\mathrm{R},k}}_{\text{local noise}}. \tag{6}$$

The pre-processing interference-plus-noise covariance matrix at receiver k is

$$\mathbf{R}_k = \sum_{\substack{q=1 \\ q \neq k}}^{K} \boldsymbol{\mathcal{T}}_{k,q}\boldsymbol{\mathcal{T}}_{k,q}^* + \sum_{m=1}^{M} \sigma_{\mathrm{X},m}^2 \boldsymbol{\mathcal{G}}_{k,m}\boldsymbol{\mathcal{G}}_{k,m}^* + \sigma_{\mathrm{R},k}^2 \mathbf{I}_{d_k}. \tag{7}$$

For notational convenience, we denote $\{\mathbf{F}\} \triangleq \{\mathbf{F}_k\}_{k=1}^{K}$, $\{\mathbf{U}\} \triangleq \{\mathbf{U}_m\}_{m=1}^{M}$ and $\{\mathbf{W}\} \triangleq \{\mathbf{W}_k\}_{k=1}^{K}$. We also denote $\boldsymbol{\mathcal{U}}_{m,k} \triangleq \mathbf{U}_m\mathbf{H}_{m,k}$ and $\boldsymbol{\mathcal{W}}_{k,m} \triangleq \mathbf{W}_k^*\mathbf{G}_{k,m}$ for $k \in \mathcal{K}$ and $m \in \mathcal{M}$. Table I summarizes the notation of equivalent channel gain matrices used in the paper for $k, q \in \mathcal{K}$ and $m \in \mathcal{M}$.

TABLE I
NOTATION OF EQUIVALENT CHANNEL MATRICES USED IN THE PAPER FOR $k, q \in \mathcal{K}$ AND $m \in \mathcal{M}$.

Equivalent channel matrix	Definition
$\boldsymbol{\mathcal{H}}_{m,k}$	$\mathbf{H}_{m,k}\mathbf{F}_k$
$\boldsymbol{\mathcal{G}}_{k,m}$	$\mathbf{G}_{k,m}\mathbf{U}_m$
$\boldsymbol{\mathcal{U}}_{m,k}$	$\mathbf{U}_m\mathbf{H}_{m,k}$
$\boldsymbol{\mathcal{W}}_{k,m}$	$\mathbf{W}_k^*\mathbf{G}_{k,m}$
$\boldsymbol{\mathcal{T}}_{k,q}$	$\sum_{m=1}^{M} \boldsymbol{\mathcal{G}}_{k,m}\boldsymbol{\mathcal{H}}_{m,q} = \sum_{m=1}^{M} \mathbf{G}_{k,m}\mathbf{U}_m\mathbf{H}_{mq}\mathbf{F}_q$

III. PROBLEM FORMULATION AND PROPOSED APPROACH

We formulate the end-to-end sum-rate maximization problem in Section III-A and propose an approach to solving it in Section III-B.

A. End-to-end Sum-Rate Maximization

For tractable analysis, we assume Gaussian signaling is used. For a given $\{\mathbf{F}\}$ and $\{\mathbf{U}\}$, the achievable rate for the k-th transmitter-receiver pair is maximized by using the linear MMSE receive filter [43]

$$\mathbf{W}_k^{\text{MMSE}} = (\boldsymbol{\mathcal{T}}_{k,k}\boldsymbol{\mathcal{T}}_{k,k}^* + \mathbf{R}_k)^{-1}\boldsymbol{\mathcal{T}}_{k,k}, \tag{8}$$

where $\mathbf{R}_k$ is given in (7). Thus, we only need to focus on the design of $\{\mathbf{F}\}$ and $\{\mathbf{U}\}$. Note that the corresponding maximum achievable rate is given by [43]

$$R_k\big(\{\mathbf{F}\}, \{\mathbf{U}\}\big) = \log_2 \det \big(\mathbf{I}_{d_k} + \boldsymbol{\mathcal{T}}_{k,k}^* \mathbf{R}_k^{-1} \boldsymbol{\mathcal{T}}_{k,k}\big). \tag{9}$$

The sum of the end-to-end achievable rates is defined as

$$R_{\text{sum}}\big(\{\mathbf{F}\}, \{\mathbf{U}\}\big) = -\sum_{k=1}^{K} \log_2 \det \Big(\mathbf{E}_k^{\text{MMSE}}(\{\mathbf{F}\}, \{\mathbf{U}\})\Big). \tag{10}$$

The end-to-end sum-rate maximization problem without power control is formulated as follows

$$(\mathcal{OP}\text{-noPC}): \min_{\{\mathbf{F}\},\{\mathbf{U}\}} \quad -R_{\text{sum}}\big(\{\mathbf{F}\}, \{\mathbf{U}\}\big)$$

$$\text{s.t.} \quad p_{\text{T},k} = p_{\text{T},k}^{\max}, k = 1, \cdots, K, \tag{11}$$

$$\sum_{m=1}^{M} p_{\text{X},m} = p_{\text{X}}^{\max}. \tag{12}$$

Remark 1: The counterpart problem with power control can be obtained by replacing the equalities in the constraints by the inequalities. We denote it as $(\mathcal{OP}\text{-PC})$. Power control may improve the end-to-end sum-rates at the expense of additional overhead because the central unit needs to inform the transmitters about both the norm and the shape of designed transmit precoders.

Remark 2: The following $(\mathbf{F}_{0,k}, \mathbf{W}_{0,k}, \mathbf{U}_{0,m})$ for $k \in \mathcal{K}$ and $m \in \mathcal{M}$ satisfies the constraints of both $(\mathcal{OP}\text{-PC})$ and $(\mathcal{OP}\text{-noPC})$

$$\mathbf{F}_{0,k} = \sqrt{\frac{p_{\text{T},k}^{\max}}{d_k}}\mathbf{I}_{N_{\text{T},k} \times d_k}, k \in \mathcal{K}, \tag{13}$$

$$\mathbf{W}_{0,k} = \sqrt{\frac{1}{d_k}}\mathbf{I}_{N_{\text{R},k} \times d_k}, k \in \mathcal{K}, \tag{14}$$

$$\mathbf{U}_{0,m} = \sqrt{\alpha p_{\text{X}}^{\max}}\mathbf{I}_{N_{\text{X},m} \times N_{\text{X},m}}, m \in \mathcal{M}, \tag{15}$$

where $\alpha = \Big(\sum_{k=1}^{K} \frac{p_{\text{T},k}^{\max}}{d_k} \sum_{m=1}^{M} \text{tr}(\mathbf{I}_{d_k \times N_{\text{T},k}}\mathbf{H}_{m,k}^*\mathbf{H}_{m,k}\mathbf{I}_{N_{\text{T},k} \times d_k}) + \sum_{m=1}^{M} N_{\text{X},m}\sigma_{\text{X},m}^2\Big)^{-1}$.

Remark 3: $(\mathcal{OP}\text{-PC})$ and $(\mathcal{OP}\text{-noPC})$ are nonconvex and NP-hard. Moreover, even the smallest configuration of the MIMO AF relay interference channel with $K = M = 2$ and $N_{\text{T}} = N_{\text{X}} = N_{\text{R}} = 2$ requires the determination of twelve complex variables for the transmit precoders and relay processing

matrices, which makes even a brute force approach challenging.

B. Proposed Approach

Instead of directly solving for the globally optimal solutions of ($\mathcal{OP}$-PC) and ($\mathcal{OP}$-noPC), we aim at finding their high-quality solutions with reasonable computational complexity. To do this, in Section III-B1 and Section III-B2, we formulate two classes of new optimizations problems that have exactly the same constraints as ($\mathcal{OP}$-PC) or ($\mathcal{OP}$-noPC) but with different objective functions.

1) Total Leakage Minimization: This section presents an approach for interference alignment in the AF relay interference channel, which is inspired by those for the single-hop interference channel in [19], [34], [35]. The underlying observation for this approach is that when interference alignment is feasible, the sum power of the interference at all the receivers, also known as the leakage, is zero. From (6), there are three groups of unwanted signals at each receiver: i) interference, ii) enhanced noise from the relays, and iii) local noise. We denote $\mathcal{I}(\{\mathbf{F}\}, \{\mathbf{U}\}, \{\mathbf{W}\})$ as the total leakage power of the AF relay interference channel. By evaluating the expectation and exploiting the independence of transmit signals $\mathbf{s}_k$ for $k \in \mathcal{K}$ and using the equality $\|\mathbf{A}\|_F^2 = \mathrm{tr}(\mathbf{A}\mathbf{A}^*)$, we obtain

$$\mathcal{I}(\{\mathbf{F}\}, \{\mathbf{U}\}, \{\mathbf{W}\}) = \sum_{k=1}^{K} \sum_{\substack{q=1 \\ q \neq k}}^{K} \mathrm{tr}(\mathbf{W}_k^* \boldsymbol{\mathcal{T}}_{k,q} \boldsymbol{\mathcal{T}}_{k,q}^* \mathbf{W}_k). \tag{16}$$

In our opinion, the high SNR regime of the relay interference channel corresponds to high transmit power at both the transmitters and the relays. As a result, in addition to eliminating completely interference, we also need to eliminate the enhanced relay noise; otherwise, the enhanced relay noise power scales with the desired signal power, preventing the system from achieving high multiplexing gain. We denote $\mathcal{N}(\{\mathbf{U}\}, \{\mathbf{W}\})$ as the sum power of enhanced noise from the relays. By evaluating the expectations and exploiting the independence of the noise vectors at the relays, we obtain

$$\mathcal{N}(\{\mathbf{U}\}, \{\mathbf{W}\}) = \sum_{k=1}^{K} \sum_{m=1}^{M} \sigma_{\mathrm{X},m}^2 \, \mathrm{tr}(\mathbf{W}_k^* \boldsymbol{\mathcal{G}}_{k,m} \boldsymbol{\mathcal{G}}_{k,m}^* \mathbf{W}_k). \tag{17}$$

Note that scaling down transmit power at either the transmitters or relays decreases the total leakage power at the receivers. For example, if $(\{(1/a)\mathbf{F}\}, \{\mathbf{U}\}, \{\mathbf{W}\})$ is used instead of $(\{\mathbf{F}\}, \{\mathbf{U}\}, \{\mathbf{W}\})$ where $\{(1/a)\mathbf{F}\} = \{(1/a)\mathbf{F}_1, \cdots, (1/a)\mathbf{F}_K\}$ and $a > 1$, then both the actual transmit power at the transmitters and the total leakage power decrease $a^2 > 1$ times. Thus, equality power constraints at the transmitters and relays are required to obtain a meaningful design problem. This means that power control should not be considered in the context of total leakage power minimization. In other words,

we do not use the total leakage minimization approach to find solutions to $(\mathcal{OP}\text{-PC})$. Also to obtain a meaningful design problem, we add the orthonormal constraints on $\mathbf{W}_k$ as $\mathbf{W}_k^* \mathbf{W}_k = \mathbf{I}_{d_k}$ for $k \in \mathcal{K}$. Without such constraints, we can always use zero matrices as the receive filters to get zero total leakage power. Consequently, to find high-quality solutions of $(\mathcal{OP}\text{-noPC})$, we propose to solve the following problem

$$(\mathcal{TL}): \quad \min_{\{\mathbf{F}\},\{\mathbf{U}\},\{\mathbf{W}\}} \quad \mathcal{I}\left(\{\mathbf{F}\},\{\mathbf{U}\},\{\mathbf{W}\}\right) + \mathcal{N}\left(\{\mathbf{U}\},\{\mathbf{W}\}\right)$$

$$\text{s.t.} \quad p_{\mathrm{T},k} = p_{\mathrm{T},k}^{\max}, k \in \mathcal{K}, \tag{18}$$

$$\sum_{m=1}^{M} p_{\mathrm{X},m} = p_{\mathrm{X}}^{\max}, \tag{19}$$

$$\mathbf{W}_k^* \mathbf{W}_k = \mathbf{I}_{d_k}, k \in \mathcal{K}. \tag{20}$$

Note that $(\mathcal{TL})$ is nonconvex and in general is NP-hard. Also, $(\mathcal{TL})$ does not take into account the desired signal power and local noise at the receivers.

Remark 4: The total leakage minimization problem formulated in [29] for an AF relay network is a simplified version of $(\mathcal{TL})$. It is assumed in [29] that the transmitters and receivers are equipped with a single antenna. Each pair is aided by a dedicated multiple-antenna AF relay. The formulation in [29] does not consider power constraints at the relays. In addition, it is assumed that there are no cross-links for the transmissions from relays to receivers, i.e. $\mathbf{G}_{k,q} = \underline{0}$ for all $k, q \in \mathcal{K}$ and $k \neq q$. As a result, for fixed $\{\mathbf{F}\}$ and $\{\mathbf{W}\}$, the algorithm in [29] can determine each $\mathbf{U}_m$ separately.

2) Sum Mean Squared Error Minimization: The section presents another approach that is based on a relationship between the achievable rates and MSE values at the receivers with Gaussian signaling [44]. This is inspired by prior work on the MIMO broadcast channel [45], MIMO interference channel [33], MIMO interference broadcast channel [46], and two-way relay channel [47], [48]. Let $\mathbf{E}_k(\{\mathbf{E}\}, \{\mathbf{U}\}, \mathbf{W}_k)$ be the MSE matrix at receiver k. After some manipulation, we obtain

$$\mathbf{E}_k(\{\mathbf{E}\}, \{\mathbf{U}\}, \mathbf{W}_k) = \mathbf{W}_k^*(\mathcal{T}_{k,k}\mathcal{T}_{k,k}^* + \mathbf{R}_k)\mathbf{W}_k - \mathbf{W}_k^*\mathcal{T}_{k,k} - \mathcal{T}_{k,k}^*\mathbf{W}_k + \mathbf{I}_{d_k}. \tag{21}$$

Note that the MSE at receiver k, defined as $MSE_k = \mathrm{tr}(\mathbf{E}_k(\{\mathbf{E}\}, \{\mathbf{U}\}, \mathbf{W}_k))$, is minimized by the linear MMSE receive filter $\mathbf{W}_k^{\mathrm{MMSE}}$. Moreover, it is well-established that [33], [44]–[49]

$$R_k\left(\{\mathbf{F}\}, \{\mathbf{U}\}\right) = -\log_2 \det(\mathbf{E}_k(\{\mathbf{F}\}, \{\mathbf{U}\}, \mathbf{W}_k^{\mathrm{MMSE}})). \tag{22}$$

We introduce auxiliary weight matrix variables $\{\mathbf{V}\} \triangleq (\mathbf{V}_1, \cdots, \mathbf{V}_K)$ that are square $(\mathbf{V}_k \in \mathbb{C}^{d_k \times d_k})$ and positive semidefinite for $k \in \mathcal{K}$. The weight matrices $\mathbf{V}_k$ are just auxiliary variables for the

optimization technique and have no actual physical meaning. Define the matrix-weighted sum of MSE values as follows

$$WMSE_{\text{sum}}(\{\mathbf{F}\}, \{\mathbf{U}\}, \{\mathbf{W}\}, \{\mathbf{V}\}) = \sum_{k=1}^{K} \left(\text{tr} \left(\mathbf{V}_k \mathbf{E}_k(\{\mathbf{F}\}, \{\mathbf{U}\}, \{\mathbf{W}\}) \right) - \log_2 \det \left(\mathbf{V}_k \right) \right). \quad (23)$$

Then, we formulate the following weighted sum-MSE minimization problem

$$(\mathcal{WMSE}\text{-noPC}): \quad \min_{\{\mathbf{F}\},\{\mathbf{U}\},\{\mathbf{W}\},\{\mathbf{V}\}} \quad WMSE_{\text{sum}}(\{\mathbf{F}\}, \{\mathbf{U}\}, \{\mathbf{W}\}, \{\mathbf{V}\})$$

$$\text{s.t.} \quad p_{\text{T},k} = p_{\text{T},k}^{\max}, k \in \mathcal{K} \quad (24)$$

$$\sum_{m=1}^{M} p_{\text{X},m} = p_{\text{X}}^{\max}. \quad (25)$$

Similarly, we formulate $(\mathcal{WMSE}\text{-PC})$ by replacing the equalities in $(\mathcal{WMSE}\text{-noPC})$. Using the same steps in [46], [49], we can show that $(\mathcal{WMSE}\text{-noPC})$ and $(\mathcal{OP}\text{-noPC})$ have exactly the same stationary points if we use the linear MMSE receivers and choose the following matrix weights

$$\mathbf{V}_k^{\text{opt}}(\{\mathbf{F}\}, \{\mathbf{U}\}, \mathbf{W}_k^{\text{MMSE}}) \quad = \quad \mathbf{E}_k^{-1}(\{\mathbf{F}\}, \{\mathbf{U}\}, \mathbf{W}_k^{\text{MMSE}}), \quad (26)$$

$$= \quad \mathbf{I}_{d_k} + \boldsymbol{\mathcal{T}}_{k,k}^* \mathbf{R}_k^{-1} \boldsymbol{\mathcal{T}}_{k,k}. \quad (27)$$

This observation is also true for $(\mathcal{WMSE}\text{-PC})$ and $(\mathcal{OP}\text{-PC})$. Thus, instead of directly solving $(\mathcal{OP}\text{-noPC})$ (or $(\mathcal{OP}\text{-PC})$), we can focus on finding high-quality solutions to its corresponding weighted sum-MSE minimization problem, which has a better-behaved objective function.

Remark 5: The matrix-weighted sum-MSE value $WMSE_{\text{sum}}(\{\mathbf{F}\}, \{\mathbf{U}\}, \{\mathbf{W}\}, \{\mathbf{V}\})$ is convex with respect to $\mathbf{F}_k$ for $k \in \mathcal{K}$ if we always choose $\mathbf{V}_k$ according to (26). Indeed, we can check that MSE_k is convex with respect to $\mathbf{F}_q$ for all $k, q \in \mathcal{K}$. By construction, $\mathbf{V}_k^{\text{opt}}(\{\mathbf{F}\}, \{\mathbf{U}\}, \mathbf{W}_k^{\text{MMSE}})$ is a Hermitian and positive semidefinite matrix for $k \in \mathcal{K}$. Then, by definition $WMSE_{\text{sum}}(\{\mathbf{F}\}, \{\mathbf{U}\}, \{\mathbf{W}\}, \{\mathbf{V}\})$ is also convex with respect to $\mathbf{F}_k$ for $k \in \mathcal{K}$.

IV. ALGORITHMS

The problems $(\mathcal{TL})$, $(\mathcal{WMSE}\text{-noPC})$, and $(\mathcal{WMSE}\text{-PC})$ formulated in Section III-B are nonconvex and in general are NP-hard. In this section, rather than attempting solving for their globally optimal solutions, we adopt an alternating minimization approach [36] to develop iterative algorithms for finding their high-quality solutions. In each iteration, we alternatively fix all but one variable and determine the remaining variable by solving a single-variable optimization problem. The optimization problem in each iteration is always feasible since it has the outcome of the previous iteration as a feasible

point. After initialization, the algorithms are repeated until a convergent point is reached. Section IV-A presents the algorithm for solving $(\mathcal{TL})$, which is denoted as Algorithm 1. Two algorithms for solving $(\mathcal{WMSE}\text{-noPC})$ and $(\mathcal{WMSE}\text{-PC})$ are presented in Section IV-B and Section IV-C. They are denoted as Algorithm 2 and Algorithm 3, respectively.

A. Algorithm for Total Leakage Minimization $(\mathcal{TL})$

There are three classes of design subproblems in Algorithm 1: i) receiver filter design, ii) relay processing matrix design, and iii) transmit precoder design.

1) Receive Filter Design for $(\mathcal{TL})$: We can rewrite the cost function as $\sum_{k=1}^{K} \operatorname{tr}(\mathbf{W}_k^* \mathbf{Z}_k \mathbf{W}_k)$, where $\mathbf{Z}_k = \sum_{\substack{q=1 \\ q \neq k}}^{K} \boldsymbol{\mathcal{T}}_{k,q} \boldsymbol{\mathcal{T}}_{k,q}^* + \sum_{m=1}^{M} \sigma_{\mathrm{X},m}^2 \boldsymbol{\mathcal{G}}_{k,m} \boldsymbol{\mathcal{G}}_{k,m}^*$. Since $\mathbf{W}_k$ for $k \in \mathcal{K}$ are decoupled in the cost function, they can be determined separately and in parallel by solving

$$(\mathcal{TL}\text{-}\mathbf{W}_k): \mathbf{W}_k = \underset{\mathbf{X} \in \mathbb{C}^{N_{\mathrm{R},k} \times d_k} : \mathbf{X}^* \mathbf{X} = \mathbf{I}_{d_k}}{\arg \min} \operatorname{tr}(\mathbf{X}^* \mathbf{Z}_k \mathbf{X}).$$

It follows from [50] that a global optimum of $(\mathcal{TL}\text{-}\mathbf{W}_k)$ is $\mathbf{W}_k = \nu_{\min}^{d_k}(\mathbf{Z}_k)$.

2) Relay Processing Matrix Design for $(\mathcal{TL})$: We focus on determining $\mathbf{U}_m$ for some $m \in \mathcal{M}$ by solving the following single-variable optimization problem $(\mathcal{TL}\text{-}\mathbf{U}_m)$

$$\min_{\mathbf{X} \in \mathbb{C}^{N_{\mathrm{X},m} \times N_{\mathrm{X},m}}} \sum_{k=1}^{K} \sum_{\substack{q=1 \\ \neq k}}^{K} \operatorname{tr}\left(\mathbf{X} \boldsymbol{\mathcal{H}}_{m,q} \boldsymbol{\mathcal{H}}_{m,q}^* \mathbf{X}^* \boldsymbol{\mathcal{W}}_{k,m}^* \boldsymbol{\mathcal{W}}_{k,m}\right) + \sigma_{\mathrm{X},m}^2 \sum_{k=1}^{K} \operatorname{tr}(\mathbf{X}^* \boldsymbol{\mathcal{W}}_{k,m}^* \boldsymbol{\mathcal{W}}_{k,m} \mathbf{X})$$

$$+ \sum_{k=1}^{K} \sum_{\substack{q=1 \\ q \neq k}}^{K} \sum_{\substack{n=1 \\ n \neq m}}^{M} \operatorname{tr}\left(\mathbf{X} \boldsymbol{\mathcal{H}}_{m,q} \boldsymbol{\mathcal{H}}_{n,q}^* \mathbf{U}_n^* \boldsymbol{\mathcal{W}}_{k,n}^* \boldsymbol{\mathcal{W}}_{k,m}\right)$$

$$+ \sum_{k=1}^{K} \sum_{\substack{q=1 \\ q \neq k}}^{K} \sum_{\substack{n=1 \\ n \neq m}}^{M} \operatorname{tr}\left(\boldsymbol{\mathcal{W}}_{k,m}^* \boldsymbol{\mathcal{W}}_{k,n} \mathbf{U}_n \boldsymbol{\mathcal{H}}_{n,q} \boldsymbol{\mathcal{H}}_{m,q}^* \mathbf{X}^*\right)$$

$$\text{s.t.} \quad \operatorname{tr}\left(\mathbf{X}\Big(\sum_{k=1}^{K} \boldsymbol{\mathcal{H}}_{m,k} \boldsymbol{\mathcal{H}}_{m,k}^* + \sigma_{\mathrm{X},m}^2 \mathbf{I}_{N_{\mathrm{X},m}}\Big)\mathbf{X}^*\right) = \eta_{\mathrm{U},m}, \tag{28}$$

where $\eta_{\mathrm{U},m} = p_{\mathrm{X}}^{\max} - \sum_{\substack{n=1 \\ n \neq m}}^{M} \sum_{k=1}^{K} \operatorname{tr}\left(\mathbf{U}_n \boldsymbol{\mathcal{H}}_{n,k} \boldsymbol{\mathcal{H}}_{n,k}^* \mathbf{U}_n^*\right) - \sum_{\substack{n=1 \\ n \neq m}}^{M} \sigma_{\mathrm{X},n}^2 \operatorname{tr}(\mathbf{U}_n \mathbf{U}_n^*)$. Because of the special form of the first term in the cost function of $(\mathcal{TL}\text{-}\mathbf{U}_m)$, it is not straightforward to use the methods for the single-hop interference channel like those in [35] to solve $(\mathcal{TL}\text{-}\mathbf{U}_m)$.

We propose to transform $(\mathcal{TL}\text{-}\mathbf{U}_m)$ into a more readily solvable form by introducing a new variable

$\mathbf{u}_m = \text{vec}(\mathbf{U}_m) \in \mathbb{C}^{N_{\text{X},m}^2 \times 1}$. We define the following matrices that are independent of $\mathbf{u}_m$

$$\mathbf{A}_{1,m} = \sum_{k=1}^{K} \left(\sum_{\substack{q=1 \\ q \neq k}}^{K} \boldsymbol{\mathcal{H}}_{m,q} \boldsymbol{\mathcal{H}}_{m,q}^* + \sigma_{\text{X},m}^2 \mathbf{I}_{N_{\text{X},m}} \right)^T \otimes \left(\boldsymbol{\mathcal{W}}_{k,m}^* \boldsymbol{\mathcal{W}}_{k,m} \right), \tag{29}$$

$$\mathbf{a}_{2,m} = \text{vec} \left(\sum_{k=1}^{K} \sum_{\substack{q=1 \\ q \neq k}}^{K} \sum_{\substack{n=1 \\ n \neq m}}^{M} \boldsymbol{\mathcal{W}}_{k,m}^* \boldsymbol{\mathcal{W}}_{k,n} \mathbf{U}_n \boldsymbol{\mathcal{H}}_{n,q} \boldsymbol{\mathcal{H}}_{m,q}^* \right), \tag{30}$$

$$\mathbf{A}_{3,m} = \left(\sum_{k=1}^{K} \boldsymbol{\mathcal{H}}_{m,k} \boldsymbol{\mathcal{H}}_{m,k}^* + \sigma_{\text{X},m}^2 \mathbf{I}_{N_{\text{X},m}} \right)^T \otimes \mathbf{I}_{N_{\text{X},m}}. \tag{31}$$

Note that with probability one, $\mathbf{A}_{3,m}$ is Hermitian and positive definite while $\mathbf{A}_{1,m}$ is Hermitian and positive semidefinite. Then, we use the following equalities, $\text{tr}(\mathbf{ABA}^*\mathbf{C}) = (\text{vec}(\mathbf{A}))^*(\mathbf{B}^T \otimes \mathbf{C}) \text{vec}(\mathbf{A})$, $\text{tr}(\mathbf{A}^*\mathbf{BA}) = \text{tr}(\mathbf{AIA}^*\mathbf{B}) = (\text{vec}(\mathbf{A}))^*(\mathbf{I} \otimes \mathbf{B}) \text{vec}(\mathbf{A})$ and $\text{tr}(\mathbf{AB}^*) = (\text{vec}(\mathbf{B}))^* \text{vec}(\mathbf{A})$ [51], to transform both the cost function and the constraint of $(\mathcal{TL}\text{-}\mathbf{U}_m)$ into quadratic expressions of $\mathbf{u}_m$. The quadratically constrained quadratic program (QCQP) for designing $\mathbf{u}_m$ is

$$(\mathcal{TL}\text{-}\mathbf{u}_m): \min_{\mathbf{x} \in \mathbb{C}^{N_{\text{X},m}^2 \times 1}} \mathbf{x}^* \mathbf{A}_{1,m} \mathbf{x} + \mathbf{a}_{2,m}^* \mathbf{x} + \mathbf{x}^* \mathbf{a}_{2,m}$$

$$\text{s.t.} \quad \mathbf{x}^* \mathbf{A}_{3,m} \mathbf{x} = \eta_{\text{U},m}. \tag{32}$$

This is a QCQP with a single equality quadratic constraint. It is nonconvex as well.

In solving $(\mathcal{TL}\text{-}\mathbf{u}_m)$, we introduce a new variable $\mathbf{Y} = \begin{pmatrix} \mathbf{u}_m \\ 1 \end{pmatrix} \begin{pmatrix} \mathbf{u}_m^* & 1 \end{pmatrix} = \begin{pmatrix} \mathbf{u}_m \mathbf{u}_m^* & \mathbf{u}_m \\ \mathbf{u}_m^* & 1 \end{pmatrix} \in$

$\mathbb{C}^{(N_{\text{X},m}^2+1) \times (N_{\text{X},m}^2+1)}$. It follows that $\mathbf{Y}$ is a rank-one Hermitian positive semidefinite matrix with bottom right entry equal to 1. We can rewrite $(\mathcal{TL}\text{-}\mathbf{u}_m)$ equivalently as

$$(\mathcal{TL}\text{-}\mathbf{u}_m\mathbf{u}_m^*): \min_{\mathbf{Y} \in \mathbb{C}^{(N_{\text{X},m}^2+1) \times (N_{\text{X},m}^2+1)}} \text{tr} \left(\begin{pmatrix} \mathbf{A}_{1,m} & \mathbf{a}_{2,m} \\ \mathbf{a}_{2,m}^* & 1 \end{pmatrix} \mathbf{Y} \right)$$

$$\text{s.t.} \quad \text{tr} \left(\begin{pmatrix} \mathbf{A}_{3,m} & \underline{0}_{N_{\text{X},m}^2 \times 1} \\ \underline{0}_{1 \times N_{\text{X},m}^2} & 1 \end{pmatrix} \mathbf{Y} \right) = \eta_{\text{U},m} + 1, \tag{33}$$

$$\text{tr} \left(\begin{pmatrix} \underline{0}_{N_{\text{X},m}^2 \times N_{\text{X},m}^2} & \underline{0}_{N_{\text{X},m}^2 \times 1} \\ \underline{0}_{1 \times N_{\text{X},m}^2} & 1 \end{pmatrix} \mathbf{Y} \right) = 1, \tag{34}$$

$$\mathbf{Y} \succeq \underline{0}, \quad \text{rank}(\mathbf{Y}) = 1. \tag{35}$$

While the cost function and all other constraints are convex, the rank constraint is nonconvex. This rank constraint is actually the main difficulty in solving $(\mathcal{TL}\text{-}\mathbf{u}_m\mathbf{u}_m^*)$. Dropping this rank constraint, however, we obtain a relaxed version of $(\mathcal{TL}\text{-}\mathbf{u}_m\mathbf{u}_m^*)$, which is a convex optimization problem and also known

as a semidefinite relaxation (SDR) of $(\mathcal{TL}\text{-}\mathbf{u}_m\mathbf{u}_m^*)$. Note that a complex-valued separable homogeneous QCQP with n constraints is guaranteed to have a global optimum with rank $r \leq \sqrt{n}$ [52]. Therefore, having $n = 1$ constraints, $(\mathcal{TL}\text{-}\mathbf{u}_m\mathbf{u}_m^*)$ is guaranteed to have a rank-one global optimum. The SDR of $(\mathcal{TL}\text{-}\mathbf{u}_m\mathbf{u}_m^*)$ can be solved, to any arbitrary accuracy, in a numerically reliable and efficient manner by readily available software packages, e.g., the convex optimization toolbox CVX [53]. It is not guaranteed, however, that solving the SDR by the available software packages provides a desired rank-one global optimum of the SDR. Fortunately, we can construct a rank-one global optimum of the SDR from the resulting general-rank global optimum using the rank-reduction procedure in [52], which is an extension of the purification technique in [54]. The key idea in each step of the procedure is to modify the eigenvalues of the general-rank global optima to remove the largest eigenvalue. Each step of the procedure gives us another global optima with the same eigenvectors but with one fewer nonzero eigenvalues. We notice that the last entry of the column vector obtained by the decomposition of the rank-one global optimum [55] may be a complex number with modulus of 1. By multiplying the resulting column vector with the conjugate of its last entry, we obtain a desired column vector in the form of $(\mathbf{u}_m \ 1)^T$, which corresponds to another rank-one global optimum of $(\mathcal{TL}\text{-}\mathbf{u}_m\mathbf{u}_m^*)$. We then use the vec^{-1} operator to get a globally optimal solution $\mathbf{U}_m$ of $(\mathcal{TL}\text{-}\mathbf{U}_m)$ from the resulting $\mathbf{u}_m$.

3) Transmit Precoder Design for $(\mathcal{TL})$: We now focus on designing $\mathbf{F}_k$ for some $k \in \mathcal{K}$ by solving the following single-variable optimization problem

$$(\mathcal{TL}\text{-}\mathbf{F}_k): \quad \min_{\mathbf{X} \in \mathbb{C}^{N_{\text{T},k} \times d_k}} \quad \text{tr}\left(\mathbf{X}^*\left(\sum_{\substack{q=1 \\ q \neq k}}^{K} \sum_{m=1}^{M} \sum_{n=1}^{M} \boldsymbol{\mathcal{U}}_{m,k}^* \boldsymbol{\mathcal{W}}_{q,m}^* \boldsymbol{\mathcal{W}}_{q,n} \boldsymbol{\mathcal{U}}_{n,k}\right)\mathbf{X}\right)$$

$$\text{s.t.} \quad \text{tr}(\mathbf{X}^*\mathbf{X}) = p_{\text{T},k} \tag{36}$$

$$\text{tr}\left(\mathbf{X}^*\left(\sum_{m=1}^{M} \boldsymbol{\mathcal{U}}_{m,k}^* \boldsymbol{\mathcal{U}}_{m,k}\right)\mathbf{X}\right) = \eta_{\text{F},k}, \tag{37}$$

where $\eta_{\text{F},k} = p_{\text{X}}^{\max} - \sum_{\substack{q=1 \\ q \neq k}}^{K} \sum_{m=1}^{M} \text{tr}\left(\mathbf{F}_q^* \boldsymbol{\mathcal{U}}_{m,q}^* \boldsymbol{\mathcal{U}}_{m,q} \mathbf{F}_q\right) - \sum_{m=1}^{M} \sigma_{\text{X},m}^2 \text{tr}\left(\mathbf{U}_m \mathbf{U}_m^*\right)$. Note that $(\mathcal{TL}\text{-}\mathbf{F}_k)$ is non-convex and in general is NP-hard. Since $(\mathcal{TL}\text{-}\mathbf{F}_k)$ has two equality constraints, the use of the Lagrange multiplier method requires a more complicated 2-D search.

Similar to Section IV-A2, we propose a method for transforming $(\mathcal{TL}\text{-}\mathbf{F}_k)$ into an equivalent optimization problem and for solving for its global optimum. We start by defining a new variable $\mathbf{f}_k = \text{vec}(\mathbf{F}_k) \in$

$\mathbb{C}^{N_{\mathrm{T},k} d_k \times 1}$. We also define the following matrices which are independent of $\mathbf{f}_k$

$$
\mathbf{B}_{1,k} = \mathbf{I}_{d_k} \otimes \left(\sum_{\substack{q=1 \\ q \neq k}}^{K} \sum_{m=1}^{M} \sum_{n=1}^{M} \boldsymbol{\mathcal{U}}_{m,k}^* \boldsymbol{\mathcal{W}}_{q,m}^* \boldsymbol{\mathcal{W}}_{q,n} \boldsymbol{\mathcal{U}}_{n,k} \right), \tag{38}
$$

$$
\mathbf{B}_{2,k} = \mathbf{I}_{d_k} \otimes \left(\sum_{m=1}^{M} \boldsymbol{\mathcal{U}}_{m,k}^* \boldsymbol{\mathcal{U}}_{m,k} \right). \tag{39}
$$

Both $\mathbf{B}_{1,k}$ and $\mathbf{B}_{2,k}$ are Hermitian positive definite matrices. Using $\mathrm{tr}(\mathbf{A}^* \mathbf{B} \mathbf{A}) = (\mathrm{vec}(\mathbf{A}))^* (\mathbf{I} \otimes \mathbf{B}) \mathrm{vec}(\mathbf{A})$ [51], we transform $(\mathcal{TL}\text{-}\mathbf{f}_k)$ into the following single-variable optimization problem

$$
(\mathcal{TL}\text{-}\mathbf{f}_k): \min_{\mathbf{x} \in \mathbb{C}^{N_{\mathrm{T},k} d_k \times 1}} \quad \mathbf{x}^* \mathbf{B}_{1,k} \mathbf{x}
$$

$$
\text{s.t.} \quad \mathbf{x}^* \mathbf{x} = p_{\mathrm{T},k}^{\max}, \tag{40}
$$

$$
\mathbf{x}^* \mathbf{B}_{2,k} \mathbf{x} = \eta_{\mathrm{F},k}. \tag{41}
$$

Note that $(\mathcal{TL}\text{-}\mathbf{f}_k)$ is a complex-valued homogeneous QCQP with two equality quadratic constraints. Nevertheless, $(\mathcal{TL}\text{-}\mathbf{f}_k)$ is still nonconvex and NP-hard [52], [56].

In solving $(\mathcal{TL}\text{-}\mathbf{f}_k)$, we introduce a new variable $\mathbf{Y} = \mathbf{x}\mathbf{x}^*$. Note that $\mathbf{Y} = \mathbf{x}\mathbf{x}^*$ requires that $\mathbf{Y}$ be a rank-one Hermitian positive semidefinite matrix. In addition, since $\mathbf{a}^* \mathbf{B} \mathbf{a} = \mathrm{tr}(\mathbf{B}\mathbf{a}\mathbf{a}^*)$ for any matrix $\mathbf{B}$ and any vector $\mathbf{a}$ [57], we obtain an equivalent optimization problem of $(\mathcal{TL}\text{-}\mathbf{f}_k)$ as follows

$$
(\mathcal{TL}\text{-}\mathbf{f}_k \mathbf{f}_k^*): \min_{\mathbf{Y} \in \mathbb{C}^{N_{\mathrm{T},k} d_k \times N_{\mathrm{T},k} d_k}} \quad \mathrm{tr}(\mathbf{B}_{1,k} \mathbf{Y})
$$

$$
\text{s.t.} \quad \mathrm{tr}(\mathbf{Y}) = p_{\mathrm{T},k}^{\max}, \tag{42}
$$

$$
\mathrm{tr}(\mathbf{B}_{2,k} \mathbf{Y}) = \eta_{\mathrm{F},k}, \tag{43}
$$

$$
\mathbf{Y} \succeq \underline{0}, \quad \mathrm{rank}(\mathbf{Y}) = 1. \tag{44}
$$

Similar to solving $(\mathcal{TL}\text{-}\mathbf{u}_m \mathbf{u}_m^*)$, we adopt the SDP method for solving $(\mathcal{TL}\text{-}\mathbf{f}_k \mathbf{f}_k^*)$. Since $(\mathcal{TL}\text{-}\mathbf{f}_k)$ is a complex-valued separable homogeneous QCQP with $n = 2$ constraints, it is guaranteed that $(\mathcal{TL}\text{-}\mathbf{f}_k \mathbf{f}_k^*)$ has a global optimum of rank $r = 1 \leq \sqrt{n}$. We can use readily available software packages, e.g., the convex optimization toolbox CVX [53], to solve for a general rank global optimum of the SDR of $(\mathcal{TL}\text{-}\mathbf{f}_k \mathbf{f}_k^*)$. Next, we can always construct a rank-one global optimum of the SDR from any of its general-rank global optimum, e.g., by using the rank reduction procedure in [52]. The decomposition of the rank-one global optimum [55] gives us the desired $\mathbf{f}_k$. Finally, we use the vec^{-1} operator to get a globally optimal solution $\mathbf{F}_k$ of $(\mathcal{TL}\text{-}\mathbf{F}_k)$ from the resulting $\mathbf{f}_k$.

B. Algorithm for Sum MSE Minimization without Power Control ($\mathcal{WMSE}$-noPC)

The design subproblems in the iterations of Algorithm 2 belong to one of the following four categories.

1) Matrix Weight Design for ($\mathcal{WMSE}$-noPC)*:* Since the matrix weights $\mathbf{V}_k^{\text{opt}}$ for $k \in \mathcal{K}$ are independent of each other, they can be updated in parallel based on (26).

2) Receive Filter Design for ($\mathcal{WMSE}$-noPC)*:* Recall that this approach requires the receivers use the linear MMSE receive filters $\mathbf{W}_k^{\text{MMSE}}$ given in (8). The receive filters can be updated in parallel.

3) Relay Processing Matrix Design for ($\mathcal{WMSE}$-noPC)*:* We focus on the design of $\mathbf{U}_m$ for some $m \in \mathcal{M}$. By substituting (7) and (21) into (23) and removing the terms independent of $\mathbf{U}_m$, we obtain the objective function of the design problem for $\mathbf{U}_m$. Let ($\mathcal{WMSE}$-noPC-$\mathbf{U}_m$) denote the design problem of $\mathbf{U}_m$. After some manipulation and using the fact that $\mathbf{V}_k$ is Hermitian, we obtain the formulation of ($\mathcal{WMSE}$-noPC-$\mathbf{U}_m$) as

$$
\begin{aligned}
\min_{\mathbf{X} \in \mathbb{C}^{N_{\text{x},m} \times N_{\text{X},m}}} \quad & \sum_{k=1}^{K} \sum_{q=1}^{K} \text{tr}\left(\mathbf{X} \mathcal{H}_{m,q} \mathcal{H}_{m,q}^* \mathbf{X}^* \mathcal{W}_{k,m}^* \mathbf{V}_k \mathcal{W}_{k,m} \right) + \sigma_{\text{X},m}^2 \sum_{k=1}^{K} \text{tr}(\mathbf{X}^* \mathcal{W}_{k,m}^* \mathbf{V}_k \mathcal{W}_{k,m} \mathbf{X}) \\
& - \sum_{k=1}^{K} \text{tr}(\mathcal{H}_{m,k} \mathbf{V}_k^* \mathcal{W}_{k,m} \mathbf{X}) + \sum_{k=1}^{K} \sum_{q=1}^{K} \sum_{\substack{n=1 \\ n \neq m}}^{M} \text{tr}(\mathcal{H}_{m,q} \mathcal{H}_{n,q}^* \mathbf{U}_n^* \mathcal{W}_{k,n}^* \mathbf{V}_k \mathcal{W}_{k,m} \mathbf{X}) \\
& - \sum_{k=1}^{K} \text{tr}(\mathbf{X}^* \mathcal{W}_{k,m}^* \mathbf{V}_k \mathcal{H}_{m,k}^*) + \sum_{k=1}^{K} \sum_{q=1}^{K} \sum_{\substack{n=1 \\ n \neq m}}^{M} \text{tr}(\mathbf{X}^* \mathcal{W}_{k,m}^* \mathbf{V}_k^* \mathcal{W}_{k,n} \mathbf{U}_n \mathcal{H}_{n,q} \mathcal{H}_{m,q}^*) \\
\text{s.t.} \quad & \text{tr}\left(\mathbf{X}^* \left(\sum_{k=1}^{K} \mathcal{H}_{m,k} \mathcal{H}_{m,k}^* + \sigma_{\text{X},m}^2 \mathbf{I}_{N_{\text{x},m}} \right) \mathbf{X} \right) = \eta_{\text{U},m}.
\end{aligned} \tag{45}
$$

Note that ($\mathcal{WMSE}$-noPC-$\mathbf{U}_m$) differs from ($\mathcal{TL}$-$\mathbf{U}_m$) mainly due to the appearance of $\mathbf{V}_k$ in the cost function. We introduce a new variable $\mathbf{u}_m = \text{vec}(\mathbf{U}_m)$ and define the following matrices

$$
\mathbf{C}_{1,m} = \sum_{k=1}^{K} \left(\sum_{q=1}^{K} \mathcal{H}_{m,q} \mathcal{H}_{m,q}^* + \sigma_{\text{X},m}^2 \mathbf{I}_{N_{\text{x},m}} \right)^T \otimes \left(\mathcal{W}_{k,m}^* \mathbf{V}_k \mathcal{W}_{k,m} \right), \tag{46}
$$

$$
\mathbf{c}_{2,m} = \text{vec}\left(- \sum_{k=1}^{K} \mathcal{W}_{k,m}^* \mathbf{V}_k \mathcal{H}_{m,k}^* + \sum_{\substack{n=1 \\ n \neq m}}^{M} \sum_{k=1}^{K} \sum_{q=1}^{K} \mathcal{W}_{k,m}^* \mathbf{V}_k \mathcal{W}_{k,n} \mathbf{U}_n \mathcal{H}_{n,q} \mathcal{H}_{m,q}^* \right). \tag{47}
$$

Using the same manipulation as in Section IV-A2 and denoting $\mathbf{C}_{3,m} = \mathbf{A}_{3,m}$, we obtain the following equivalent optimization problem

$$
(\mathcal{WMSE}\text{-noPC-}\mathbf{u}_m): \quad \min_{\mathbf{x} \in \mathbb{C}^{N_{\text{X},m}^2 \times 1}} \quad \mathbf{x}^* \mathbf{C}_{1,m} \mathbf{x} + \mathbf{c}_{2,m}^* \mathbf{x} + \mathbf{x}^* \mathbf{c}_{2,m}
$$

$$
\text{s.t.} \quad \mathbf{x}^* \mathbf{C}_{3,m} \mathbf{x} = \eta_{\text{U},m}. \tag{48}
$$

Note that ($\mathcal{WMSE}$-noPC-$\mathbf{u}_m$) has exactly the same form as ($\mathcal{TL}$-$\mathbf{u}_m$), thus we can apply the same

method used for solving ($\mathcal{TL}$-$\mathbf{u}_m$) to find a globally optimal solution $\mathbf{u}_m$ of ($\mathcal{WMSE}$-noPC-$\mathbf{u}_m$). We then use the vec^{-1} operator to get a globally optimal solution $\mathbf{U}_m$ of ($\mathcal{WMSE}$-noPC-$\mathbf{U}_m$) from the resulting $\mathbf{u}_m$.

4) Transmit Precoder Design for ($\mathcal{WMSE}$-noPC)*:* We define the following matrices for the design of $\mathbf{F}_k$ for some $k \in \mathcal{K}$

$$\mathbf{D}_{1,k} = \sum_{q=1}^{K}\sum_{m=1}^{M}\sum_{n=1}^{M}\mathcal{U}_{m,k}^{*}\mathcal{W}_{q,m}^{*}\mathbf{V}_q\mathcal{W}_{q,n}\mathcal{U}_{n,k}, \tag{49}$$

$$\mathbf{D}_{2,k} = \sum_{m=1}^{M}\mathcal{U}_{m,k}^{*}\mathcal{W}_{k,m}^{*}\mathbf{V}_k^{*}, \tag{50}$$

$$\mathbf{D}_{3,k} = \sum_{m=1}^{M}\mathcal{U}_{m,k}^{*}\mathcal{U}_{m,k}. \tag{51}$$

After some manipulation, we obtain the following single-variable optimization problem

$$(\mathcal{WMSE}\text{-noPC-}\mathbf{F}_k): \min_{\mathbf{X}\in\mathbb{C}^{N_{\mathrm{T},k}\times d_k}} \quad \text{tr}(\mathbf{X}^{*}\mathbf{D}_{1,k}\mathbf{X}) - \text{tr}(\mathbf{D}_{2,k}^{*}\mathbf{X}) - \text{tr}(\mathbf{D}_{2,k}\mathbf{X}^{*})$$

$$\text{s.t.} \quad \text{tr}(\mathbf{X}^{*}\mathbf{X}) = p_{\mathrm{T}}^{\max}, \tag{52}$$

$$\text{tr}(\mathbf{X}^{*}\mathbf{D}_{3,k}\mathbf{X}) = \eta_{\mathrm{T},k}. \tag{53}$$

Recall that we define $\mathbf{f}_k = \text{vec}(\mathbf{F}_k)$. We introduce a new variable $\mathbf{Y} = \begin{pmatrix}\mathbf{f}_k \\ 1\end{pmatrix}\begin{pmatrix}\mathbf{f}_k^{*} & 1\end{pmatrix} = \begin{pmatrix}\mathbf{f}_k\mathbf{f}_k^{*} & \mathbf{f}_k \\ \mathbf{f}_k^{*} & 1\end{pmatrix}$.

It follows that $\mathbf{Y}$ is a rank-one Hermitian positive semidefinite matrix with the bottom right entry equal to 1. Then, we transform ($\mathcal{WMSE}$-noPC-$\mathbf{F}_k$) equivalently into the following problem

$$(\mathcal{WMSE}\text{-noPC-}\mathbf{f}_k\mathbf{f}_k^{*}): \min_{\mathbf{Y}\in\mathbb{C}^{(N_{\mathrm{T},k}d_k+1)\times(N_{\mathrm{T},k}d_k+1)}} \quad \text{tr}\left(\begin{pmatrix}\mathbf{I}_{d_k}\otimes\mathbf{D}_{1,k} & -\text{vec}(\mathbf{D}_{2,k}) \\ -(\text{vec}(\mathbf{D}_{2,k}))^{*} & 1\end{pmatrix}\mathbf{Y}\right)$$

$$\text{s.t.} \quad \text{tr}(\mathbf{Y}) = p_{\mathrm{T}}^{\max} + 1, \tag{54}$$

$$\text{tr}\left(\begin{pmatrix}\mathbf{I}_{d_k}\otimes\mathbf{D}_{3,k} & \underline{0}_{N_{\mathrm{T},k}d_k\times 1} \\ \underline{0}_{1\times N_{\mathrm{T},k}d_k} & 1\end{pmatrix}\mathbf{Y}\right) = \eta_{\mathrm{T},k} + 1, \tag{55}$$

$$\text{tr}\left(\begin{pmatrix}\underline{0}_{N_{\mathrm{T},k}d_k\times N_{\mathrm{T},k}d_k} & \underline{0}_{N_{\mathrm{T},k}d_k\times 1} \\ \underline{0}_{1\times N_{\mathrm{T},k}d_k} & 1\end{pmatrix}\mathbf{Y}\right) = 1, \tag{56}$$

$$\mathbf{Y} \succeq \underline{0}, \quad \text{rank}(\mathbf{Y}) = 1. \tag{57}$$

Note that ($\mathcal{WMSE}$-noPC-$\mathbf{f}_k\mathbf{f}_k^{*}$) has the same form as ($\mathcal{TL}$-$\mathbf{u}_m\mathbf{u}_m^{*}$) but with one more constraint. Since ($\mathcal{WMSE}$-noPC-$\mathbf{f}_k\mathbf{f}_k^{*}$) has $n = 3$ constraints (excluding the rank-one constraint), its SDR obtained

by relaxing the rank-one constraint is exact [52]. Thus, we can use the same steps as those in solving $(\mathcal{TL}\text{-}\mathbf{u}_m\mathbf{u}_m^*)$ to find the desired column vector $\mathbf{f}_k$ corresponding to a rank-one global optimum of $(\mathcal{WMSE}\text{-noPC-}\mathbf{f}_k\mathbf{f}_k^*)$. We then use the vec^{-1} operator to get a globally optimal solution $\mathbf{F}_k$ of $(\mathcal{WMSE}\text{-noPC-}\mathbf{F}_k)$ from the resulting $\mathbf{f}_k$.

C. Algorithm for Sum MSE Minimization with Power Control $(\mathcal{WMSE}\text{-PC})$

In this section, we discuss briefly how to solve $(\mathcal{WMSE}\text{-PC})$. Using the same steps as in Section IV-B, we can develop Algorithm 3 for finding high-quality solutions of $(\mathcal{WMSE}\text{-PC})$. The details of Algorithm 3 are provided in [38], thus here we only compare and contrast the steps of Algorithm 3 and those of Algorithm 2. First, the matrix weight and receive filter designs for $(\mathcal{WMSE}\text{-PC})$ are exactly the same as those for $(\mathcal{WMSE}\text{-noPC})$. Second, the relay processing matrix design for $(\mathcal{WMSE}\text{-PC})$ can be solved by the Lagrangian multiplier method with the only difference is that the multiplier must be nonnegative. Finally, the optimization problem for the transmit precoder design for $(\mathcal{WMSE}\text{-PC})$ is obtained by replacing the equality constraints in $(\mathcal{WMSE}\text{-noPC-}\mathbf{F}_k)$ by the corresponding inequality constraints. Fortunately, the resulting optimization problem is convex with respect to $\mathbf{F}_k$. In particular, it follows from Remark 5 that the objective function of the problem $(\mathcal{WMSE}\text{-PC-}\mathcal{F}_k)$ is convex with respect to $\mathbf{F}_k$. In addition, since $\mathbf{D}_{3,k}$ is a Hermitian and positive semidefinite matrix, then we can easily check that the constraints of the resulting problem are also convex with respect to $\mathbf{F}_k$. Thus, any available software package for convex optimization could be used to solve for its unique global optimum $\mathbf{F}_k$.

D. Discussion

In this section, we discuss the proposed algorithms in the following aspects: i) the convergence, ii) the quality of the solution, and iii) the assumption on power constraints at the relays.

In terms of convergence, we are able to find a global optimum of the single-variables minimization problem in each iteration. Thus, the cost function of the original multi-variable optimization problem is non increasing after each iteration [36], [37]. This guarantees that all the proposed algorithms are convergent. Note that the authors of [49] adopt the alternating minimization to develop an iterative algorithm for solving a weighted sum-MSE minimization problem for the single-hop MIMO interference broadcast channel. That optimization problem has a differentiable objective function and a set of separable constraints in the main variables. Using the results from the general optimization [58], the authors of [49]

are able to claim that their proposed alternating minimization algorithm converges to a stationary point of the corresponding weighted sum-MSE minimization problem, which is also a stationary point of the associated sum-rate maximization problem Nevertheless, the optimization problems in the paper, $(\mathcal{TL})$, $(\mathcal{WMSE}\text{-noPC})$, and $(\mathcal{WMSE}\text{-PC})$, have non-separable constraints due to the impact of transmitter precoders on the transmit power constraints at the relays. Thus, we are currently unable to make any strong claim about whether or not our proposed iterative algorithms always converge to a stationary point of the corresponding optimization problems.

All the proposed algorithms are not guaranteed to reach a global optimum of the corresponding multi-variable optimization problem. The quality of the resulting solution depends on the initialization. One way to improve the performance of the proposed algorithms is to use multiple initializations, selecting the one with the best performance at the expenses of running time.

The proposed algorithms in the current form are applicable only under the assumption of sum-power constraint at the relays. If individual power constraints at the relays are considered, we must formulate the corresponding optimization problems. For example, as in Section III, to find high-quality solutions of the sum-rate maximization problem with individual power inequality constraints at the relays, we can formulate the following weighted sum-MSE minimization problem

$$(\mathcal{WMSE}\text{-PC-ind}): \min_{\{\mathbf{F}\},\{\mathbf{U}\},\{\mathbf{W}\},\{\mathbf{V}\}} \quad WMSE_{\mathrm{sum}}\big(\{\mathbf{F}\},\{\mathbf{U}\},\{\mathbf{W}\},\{\mathbf{V}\}\big)$$

$$\text{s.t.} \quad p_{\mathrm{T},k} \leq p_{\mathrm{T},k}^{\max}, k = 1,\cdots,K, \tag{58}$$

$$p_{\mathrm{X},m} \leq p_{\mathrm{X},m}^{\max}, m = 1,\cdots,M. \tag{59}$$

In some cases, we can use the same steps as in the previous sections to develop new alternative minimization based algorithms for solve the counterpart problems with individual power constraints at the relays, like $(\mathcal{WMSE}\text{-PC-ind})$. Note that the main difference between the sum power constraint case and the individual power constraint case is the extra constraints in the transmit precoder design. Specifically, with per-relay power constraints, the number of quadratic constraints of the resulting QCQP for the transmit precoder design is $(M + 1)$ instead of 2 as with the sum-power constraint. The extra constraints make it impossible to use the same SDP method to a globally optimal solution of the transmit precoder design problems for individual relay power constraints when there are more than two relays, i.e., $M \geq 3$. Developing new methods to solve the counterpart problems to $(\mathcal{TL})$ or $(\mathcal{WMSE}\text{-noPC})$ when $M \geq 3$ is left for future work. When power control is considered, however, we can still use the

same steps as in the previous sections to solve for high-quality solutions to ($\mathcal{WMSE}$-PC-ind). Indeed, with power control, the single-variable optimization problems for designing the relay processing matrices and transmit precoders in solving ($\mathcal{WMSE}$-PC-ind) are convex [38]. Thus, we are always able to find their global optimum.

V. SIMULATIONS

This section presents Monte Carlo simulation results to investigate the average end-to-end sum-rate performance and to gain insights into the achieved multiplexing gains of the proposed algorithms. We consider only symmetric systems, which are denoted as $(N_{\mathrm{R}} \times N_{\mathrm{T}}, d)^K + N_{\mathrm{X}}^M$, where $N_{\mathrm{R},k} = N_{\mathrm{R}}$, $N_{\mathrm{T},k} = N_{\mathrm{T}}$, $d_k = d$ and $N_{\mathrm{X},m} = N_{\mathrm{X}}$ for $k \in \mathcal{K}, m \in \mathcal{M}$. The power values are normalized such that $\sigma_{\mathrm{R},k} = \sigma_{\mathrm{X},m} = 1$, $p_{\mathrm{T},k}^{\max} = P$, $p_{\mathrm{X},m}^{\max} = P$, and $p_{\mathrm{X}}^{\max} = MP$ for $k \in \mathcal{K}, m \in \mathcal{M}$. The channel realizations are flat in time and frequency. The channel coefficients are generated as i.i.d. zero-mean unit-variance complex Gaussian random variables. No path loss is assumed in the simulations, thus the average power of all cross-links on the same hop is the same. The plots are produced by averaging over 1000 random channel realizations. For each channel realization, the initial transceivers are chosen randomly subject to the power constraints at the transmitters and relays. The same initializations are used where applicable. Each iteration updates either one transmitter or one relay and then all the receive filters. The same order of relays or transmitters selected for updating is used where applicable, for example, in the comparison of the proposed algorithms. We use the CVX toolbox [53] to solve convex problems.

For comparison, we consider the dedicated DF relay interference channel where one DF relay is dedicated to aid one and only one transmitter-receiver pair, i.e., $K = M$. Using equal time-sharing, the end-to-end achievable rate of a pair is defined as half of the minimum between the achievable rate from the transmitter to the associate DF relay and that from the relay to the receiver. We are interested only in the performance of DF relays when spatial interference alignment strategies, like those in [34], [49], are applied on two hops. Although other interference alignment techniques, like asymmetric complex signaling [59], may improve the performance of DF relays, their impacts on DF relays are left for future work. Based on [12], we derive an upper-bound on the achievable end-to-end multiplexing gain of the dedicated DF relay interference channel $(N_{\mathrm{R}} \times N_{\mathrm{T}}, d)^K + N_{\mathrm{X}}^K$ as $0.5 * \min \left\{ \left\lfloor \frac{K(N_{\mathrm{X}} + N_{\mathrm{T}})}{K+1} \right\rfloor, \left\lfloor \frac{K(N_{\mathrm{R}} + N_{\mathrm{X}})}{K+1} \right\rfloor \right\}$. Individual power constraints at the relays are considered in the comparison of AF relays and DF relays. We also assume the transmit power at a transmitter or a relay in DF relay systems is equal to P.

1) Convergence: Fig. 2 illustrates the convergence behavior of the proposed algorithms. Fig. 2(a) provides the analysis of the sum power of post-processed leakage signals of Algorithm 1 for a random channel realization of the $(4 \times 4, 2)^3 + 4^3$ system. We observe that the sum power of leakage signals decreases monotonically over iterations. Interestingly, the interference and the enhanced relay noise change their roles during the process of Algorithm 1. The interference is dominant at the beginning, however, it can be aligned and then cancelled quickly in a few iterations. After this point, the enhanced relay noise becomes dominant - its sum power is thousands times larger than the interference sum power. Unfortunately, given that many spatial dimensions have been devoted to deal with interference, it becomes challenging for Algorithm 1 to align and cancel the enhanced relay noise power. Intuitively, the enhanced relay noise can be thought of as a source of single-hop interference from "virtual uncoordinated relays" that impacts directly the receivers. Thus, we need to take into account both the interference and enhanced relay noise in the design of interference alignment strategies for the AF relay interference channel. Fig. 2(b) provides the values of $WMSE_{\text{sum}}$ achieved by Algorithm 2 and by Algorithm 3 over iterations for a channel realization of the $(2 \times 4, 1)^4 + 2^4$ system. We observe that $WMSE_{\text{sum}}$ values for both algorithms are non increasing over iterations. Although the convergence speeds of the proposed algorithms are quite fast for these configurations, they might be slow for networks with large values of K or d.
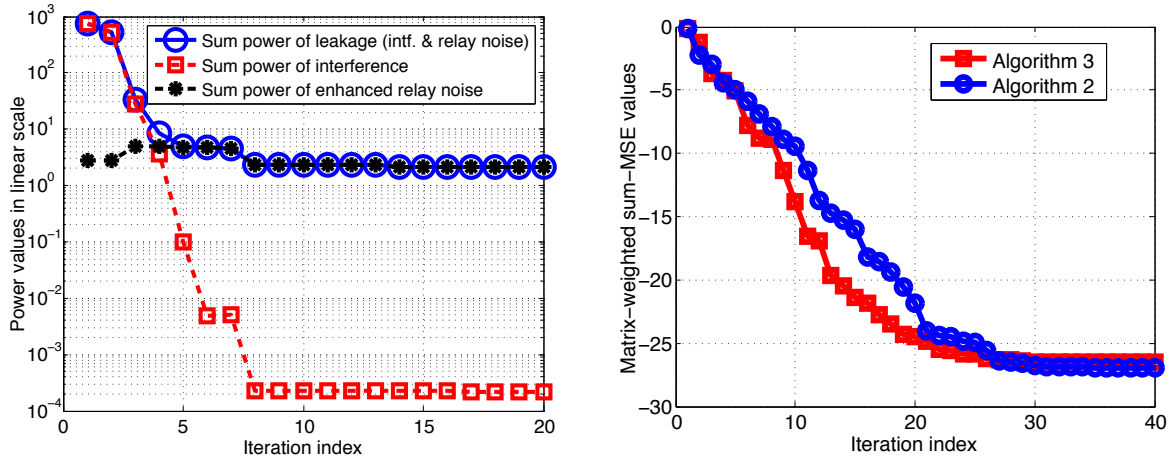
(a) Total leakage power at the receivers over iterations of Algorithm 1 for a channel realization of $(4 \times 4, 2)^3 + 4^3$.

(b) Matrix-weighted sum-MSE values over iterations of Algorithm 2 and Algorithm 3 for a channel realization of $(2 \times 4, 1)^4 + 2^4$.

Fig. 2.   Convergence behavior of the proposed algorithms.

2) Comparison of the Proposed Algorithms: In this experiment, we simulate the average achievable end-to-end sum-rates for the $(2 \times 4, 1)^4 + 2^4$ system under the sum power constraint at the relays as shown
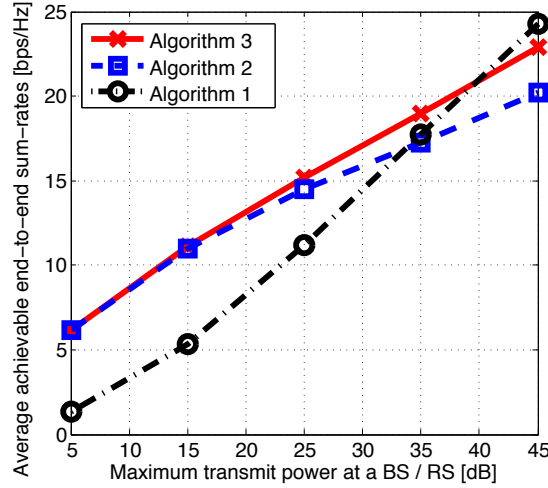
Fig. 3. Comparison of the average achievable end-to-end sum-rates of the proposed algorithms for the $(2 \times 4, 1)^4 + 2^4$ system.

in Fig. 3. We consider a sum relay power constraint. Thanks to power control, Algorithm 3 outperforms Algorithm 2 in this experiment. Both Algorithm 2 and Algorithm 3 outperform Algorithm 1 at low-to-medium SNR values because Algorithm 1 does not take into account the desired signal and noise at the receivers while the other do. Interestingly, at high SNR values, Algorithm 1 outperforms both Algorithm 2 and Algorithm 3. Especially, Algorithm 1 can achieve a higher multiplexing gain than do the other. Zooming in on per-user achievable end-to-end rates, we find that for Algorithm 2 and Algorithm 3, some users have much smaller rates than do the others; they even turn off some data streams. This unfairness limits the maximum end-to-end multiplexing gains achievable by the two algorithms. Thus, Algorithm 1 is more suitable than the others for investigating the maximum achievable end-to-end multiplexing gains of MIMO AF relay networks.

3) Sum Power Constraints vs. Individual Power Constraints at Relays: In the previous experiments, we consider sum-power constraints at the relays. In this experiment, we consider the impacts of individual power constraints. Note that any feasible point satisfies the individual power constraints at the relays also satisfies the corresponding sum-power constraint. Based on the discussion in Section IV-D, we focus on the case where power control is considered as it allows for the use of any number of relays. Fig. 4 shows the achievable end-to-end sum-rates as functions of the transmit power at a base station or a relay for both types of power constraints at the relays for the following three systems: i) $(4 \times 4, 2)^3 + 4^3$, ii) $(2 \times 2, 1)^4 + 2^4$, and iii) $(1 \times 1, 1)^3 + 2^3$. We observe that Algorithm 3 for the sum-power constraint case slightly outperforms its counterpart algorithm for the individual power constraint case in terms of

maximizing average achievable end-to-end sum-rates. This gain is due to having more freedom in power allocation in the sum-power constraint case as relays may transmit at a higher value than the maximum transmit power at a relay in the individual power constraint case. This means that extra constraints added by the individual power constraints at the relays have little impact on the end-to-end sum-rate performance of the proposed algorithms. In the following experiments, we use only the counterpart version of Algorithm 3 that is designed for individual relay power constraints, which we refer to as 'modified Algorithm 3'.
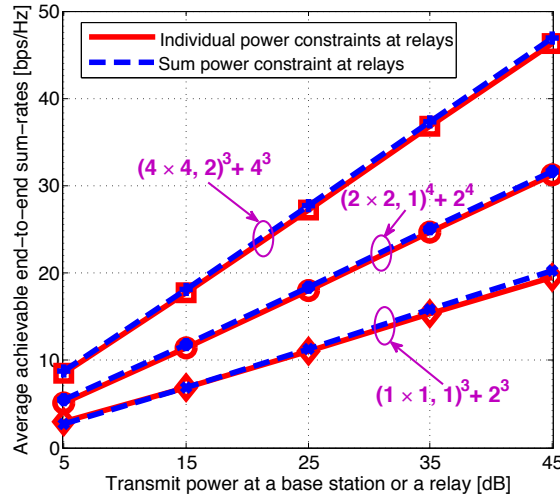
Fig. 4. Achievable end-to-end sum-rates with both types of power constraints at the relays for the $(1 \times 1, 1)^3 + 2^3$, $(2 \times 2, 1)^4 + 2^4$, and $(4 \times 4, 2)^3 + 4^3$ systems.

4) Comparison with Existing Strategies: In these experiments, we simulate several existing transceiver design strategies for the relay interference channel. For fair comparison, in this experiment and the remaining experiments, we consider the individual relay power constraints. Specifically, we simulate two strategies for the AF relay case. One is the AF TDMA distributed beamforming (BF), where all the relays help only one transmitter-receiver pair at a time (which is an extension of the design in [42] for multiple-antenna receivers). Another is the dedicated relay BF where each AF relay is devoted to aiding one and only one transmitter-receiver pair. This means that interference is ignored and we apply the joint source-relay design in [60], [61] independently for the two-hop channels from the transmitters to their associated receivers. We also three strategies for the DF relay case that correspond to independent applications of single-hop strategies on two hops. The single-hop strategies include the following: i) selfish (SF) beamforming (i.e., each transmitter aims at maximizing the achievable rate to its associated receiver), ii) interference alignment strategy based on total leakage (TL) minimization [34], and iii) the

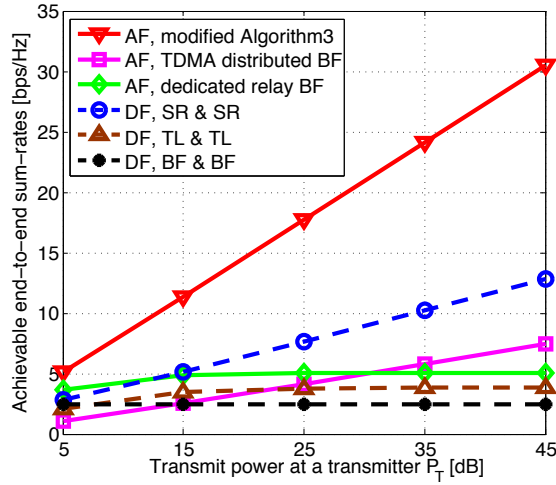iteratively weighted MSE sum-rate (SR) maximization strategy [49].

Fig. 5. Achievable end-to-end sum-rates for the $(2 \times 2, 1)^4 + 2^4$ system.

Fig. 5 shows the results for the $(2 \times 2, 1)^4 + 2^4$ system. Recall that Algorithm 3 outperforms all the other in all regions. It achieves an end-to-end multiplexing gain of 2 (which is equal to half of the total number of data streams). Note that we do not claim that this is the maximum degrees of freedom of this system. More complicated designs, for example those that can take advantage of symbol extensions [26], may achieve higher end-to-end multiplexing gains. Unaware of interference, the dedicated relay strategies for both AF relay and DF relay cases achieve zero multiplexing gains. While the multiplexing gain achieved by the DF TL & TL strategy is zero, that by the DF SR & SR strategy is nonzero. The reason is that interference alignment is not feasible for the configuration on the two hops, interference cannot be completely eliminated using the TL algorithm. Although the SR algorithm is able to turn off some data streams, one data stream on each hop in this case, to make interference alignment feasible. Note, however, that it may turn off data streams of different pairs on two hops. Thus, on average the DF SR & SR strategy achieves an end-to-end multiplexing gain less than 1.5 (half of the number of remaining data streams when interference alignment is feasible). Finally, thanks to orthogonalization transmission, the AF TDMA distributed BF can achieve an end-to-end multiplexing gain of 0.5.

Another experiment focuses on comparing Algorithm 3 with the the minimum-SINR maximization algorithm in [28]. Note that the algorithm in [28] is applicable only for single-antenna receivers. Fig. 6 shows the end-to-end achievable sum-rates of Algorithm 3 and the minimum-SINR maximization algorithm for two configurations $(1 \times 1, 1)^4 + 2^4$ and $(1 \times 1, 1)^3 + 2^2$. We observe that the end-to-end
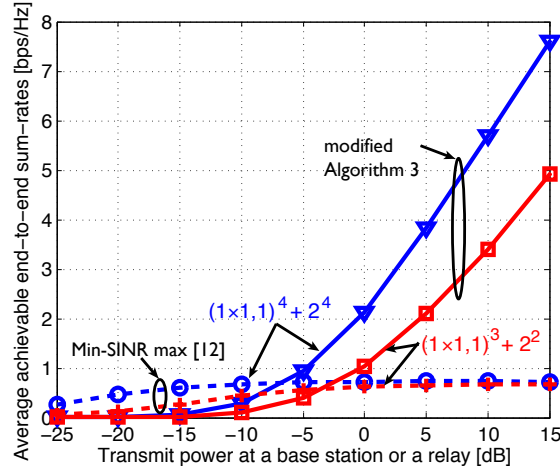
Fig. 6. Comparison of achievable end-to-end sum-rates of Algorithm 3 (shown by the solid lines) and the minimum-SINR maximization algorithm in [28] (shown by the dashed lines) for two configurations $(1 \times 1, 1)^4 + 2^4$ and $(1 \times 1, 1)^3 + 2^2$.

sum-rate performance of the minimum-SINR maximization algorithm increases at low transmit power (i.e., in the noise-limited regime) and saturates at high transmit power (i.e., the interference-limited regime). Thus, the algorithm achieves a end-to-end multiplexing gain of zero. This is reasonable since it is not designed specifically for interference management. Thanks to its capability of interference management, Algorithm 3 still achieves non-zero end-to-end multiplexing gains and provides large end-to-end sum-rate gains over the minimum-SINR maximization algorithm in the interference-limited regime. Our algorithm, however, performs worse than the the minimum-SINR maximization algorithm in the noise-limited regime where interference becomes a negligible issue.

5) Maximum Achievable Multiplexing Gains: We fix $N_{\mathrm{R}} = N_{\mathrm{T}} = 2$ and $d = 1$. Fig. 7 shows the achievable end-to-end multiplexing gains achieved by using Algorithm 1 as a function of K for $N_{\mathrm{X}} = 3$ and $N_{\mathrm{X}} = 5$ for AF relays, DF relays, and direct transmission. We notice that with these values of N_{X} and when K is small, due to the half-duplex loss, both the AF relay and DF relay cases achieve lower multiplexing gains than the direct transmission. While the DF relay case cannot outperform the direct transmission, the AF relay case can achieve higher multiplexing gains when there are more than 6 users. Thus, we can claim that AF relays help increase the achievable end-to-end multiplexing gains of interference channels. In addition, we observe that there exist upper-bounds on the achievable end-to-end multiplexing gains for all the simulated cases - AF relays, DF relays, and direct transmission. Theoretical investigation of the upper-bounds is left for future work.
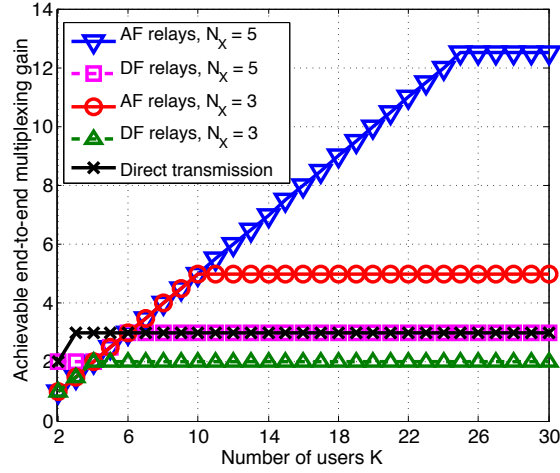
Fig. 7. Achievable end-to-end multiplexing gains as functions of K for the $(2 \times 2, 1)^K + N_X^K$ systems.
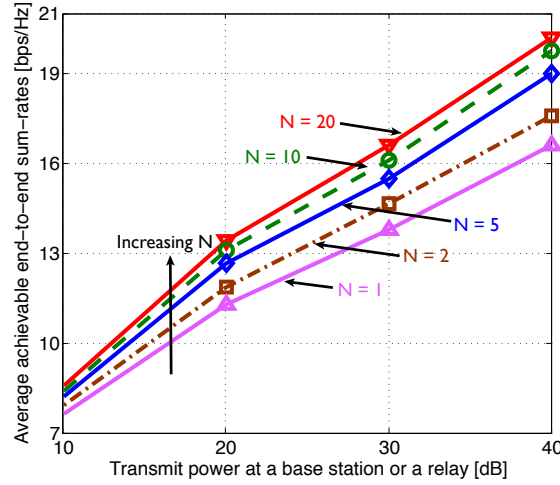

Fig. 8. Achievable end-to-end sum-rates of the opportunistic approach for the $(2 \times 2, 1)^4 + 2^4$ system with $N = \{1, 2, 5, 10, 20\}$.

6) Opportunistic Approach: The end-to-end sum-rate performance of the stationary points found by Algorithm 2 and Algorithm 3 depend significantly on the initializations. The opportunistic approach proposes to use multiple initializations and then chooses the one with the highest end-to-end sum-rates. Let N denote the number of random initializations. Fig. 8 shows the average end-to-end sum-rates for several values of N achieved by Algorithm 3 in the $(2 \times 2, 1)^4 + 2^4$ system. For this setting, at a transmit power of 30dB, the gain provided by the opportunistic approach over the non-opportunistic approach is 6.4% for $N = 2$, 13.2% for $N = 5$, 16.9% for $N = 10$, and 20.6% for $N = 20$. Note that the higher the value of N, the larger the average achievable end-to-end sum-rates. Also, the additional gains obtained by using an extra random initialization decreases in N. Nevertheless, the benefits of this opportunistic approach come at the expense of longer running time.

VI. Conclusions and Future Work

We developed three cooperative algorithms for joint designs of the transmitters, relays, and receivers of the MIMO AF relay interference channel. Algorithm 1 aims at minimizing the sum power of the interference signals and the enhanced noise from the relays. Based on a relationship between MSE and mutual information, Algorithm 2 (Algorithm 3) is able to find a stationary point of the end-to-end sum-rate maximization problems with equality (inequality) power constraints. Simulations show that thanks to the consideration of the desired signal power and the noise power at the receivers, Algorithm 2 and Algorithm 3 outperform Algorithm 1 at low-to-medium SNR. Nevertheless, they perform worse than Algorithm 1 at high SNR due to unfairness in rate allocation among users. The multiplexing gains achievable by the proposed algorithms provide lower bounds on the total number of degrees of freedom in MIMO AF relay networks, which remains unknown. Also, the use of AF relays results in higher end-to-end multiplexing gains than both the use of DF relays and the direct transmission.

A major limitation of our algorithms is that global CSI is required to implement them in their present form. Naturally this is challenging to achieve in a distributed system. We believe the results are still valuable, however, because they provide a benchmark for developing algorithms that relax the global CSI assumptions. Future work should focus on developing cooperative algorithms that require less overhead, have faster convergence speed, allow for lower implementation complexity, and account for channel estimation error

References

[1] O. Simeone, O. Somekh, Y. Bar-Ness, H. V. Poor, and S. Shamai, "Capacity of linear two-hop mesh networks with rate splitting, decode-and-forward relaying and cooperation," in *Proc. of Allerton Conf. Commun., Control, Computing*, Monticello, IL, Sep. 2007.

[2] P. S. C. Thejaswi, A. Bennatan, J. Zhang, R. Calderbank, and D. Cochran, "Rate-achievability strategies for two-hop interference flows," in *Proc. of Allerton Conf. Commun. Control Computing*, Montecillo, IL, Sep. 2008, pp. 1432–1439.

[3] Y. Cao and B. Chen, "Capacity bounds for two-hop interference networks," in *Proc. of Allerton Conf. Commun., Control, Computing*, Montecillo, IL, Sep.-Oct. 2009, pp. 272–279.

[4] M. A. Torabi and J.-F. Frigon, "A decomposition approach to MIMO interference relay networks," in *Proc. of IEEE Global Telecommun. Conf.*, New Orleans, LA, Nov.-Dec. 2008, pp. 1–6.

[5] 3GPP TR 36.814, "Further enhancements for E-UTRA physical layer aspects."

[6] IEEE 802.16m Broadband Wireless Access Working Group, "IEEE 802.16m system requirement."

[7] Y. Zhu and H. Zheng, "Understanding the impact of interference on collaborative relays," *IEEE Trans. Mobile Computing*, vol. 7, no. 6, pp. 724–736, Jun. 2008.

[8] S. W. Peters, A. Y. Panah, K. T. Truong, and R. W. Heath, Jr., "Relay architectures for 3GPP LTE-Advanced," *EURASIP J. Wireless Commun. Networking*, vol. 2009, 2009.

[9] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the K-user interference channel," *IEEE Trans. Info. Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.

[10] T. Gou and S. A. Jafar, "Degrees of freedom of the K user M x N MIMO interference channel," *IEEE Trans. Info. Theory*, vol. 56, no. 12, pp. 6040–6057, Dec. 2010.

[11] C. M. Yetis, T. Gou, S. A. Jafar, and A. H. Kayran, "On feasibility of interference alignment in MIMO interference networks," *IEEE Trans. Signal Processing*, vol. 58, no. 9, pp. 4771–4782, Sep. 2010.

[12] M. Razaviyayn, G. Lyubeznik, and Z.-Q. Luo, "On the degrees of freedom achievable through interference alignment in a MIMO interference channel," *IEEE Trans. Signal Processing*, vol. 60, no. 2, pp. 812–821, Feb. 2012.

[13] S. A. Jafar, *Interference Alignment - A new Look at Signal Dimensions in a Communication Network*. Now Publishers Inc., 2011, vol. 7, no. 1, pp. 1–134.

[14] H. Ghozlan, Y. Mohasseb, G. Kramer, and H. El Gamal, "The MIMO wireless switch: Relaying can increase the multiplexing gain," in *Proc. of IEEE Int. Symp. Info. Theory*, Jul. 2009, pp. 1448–1452.

[15] S.-W. Jeon, S.-Y. Chung, and S. A. Jafar, "Degrees of freedom region of a class of multisource Gaussian relay networks," *IEEE Trans. Info. Theory*, vol. 57, no. 5, pp. 3032–3044, May 2011.

[16] A. Host-Madsen and A. Nosratinia, "The multiplexing gain of wireless networks," in *Proc. of IEEE Int. Symp. Info. Theory*, Sep. 2005, pp. 2065–2069.

[17] S. Berger, M. Kuhn, A. Wittneben, T. Unger, and A. Klein, "Recent advances in amplify-and-forward two-hop relaying," *IEEE Commun. Mag.*, vol. 47, no. 7, pp. 50–56, july 2009.

[18] V. R. Cadambe and S. A. Jafar, "Degrees of freedom of wireless networks with relays, feedback, cooperation, and full duplex operation," *IEEE Trans. Info. Theory*, vol. 55, no. 5, pp. 2334–2344, May 2009.

[19] K. Gomadam, V. R. Cadambe, and S. A. Jafar, "A distributed numerical approach to interference alignment and applications to wireless interference networks," *IEEE Trans. Info. Theory*, vol. 57, no. 6, pp. 3309–3322, Jun. 2011.

[20] T. Abe, H. Shi, T. Asai, and H. Yoshino, "Relay techniques for MIMO wireless networks with multiple source and destination pairs," *EURASIP J. Wireless Commun. Networking*, vol. 2006, 2006.

[21] V. I. Morgenshtern and H. Bolcskei, "Crystallization in large wireless networks," *IEEE Trans. Info. Theory*, vol. 53, no. 10, pp. 3319–3349, Oct. 2007.

[22] B. Rankov and A. Wittneben, "Spectral efficient protocols for half-duplex fading relay channels," *IEEE J. Select. Areas Commun.*, vol. 25, no. 2, pp. 379–389, Feb. 2007.

[23] B. Nourani, S. A. Motahari, and A. K. Khandani, "Relay-aided interference alignment for the quasi-static interference channel," in *Proc. of IEEE Int. Symp. Info. Theory*, Austin, TX, Jun. 2010, pp. 405–409.

[24] S. Chen and R. S. Cheng, "Achieve the degrees of freedom of K-user MIMO interference channel with a MIMO relay," in *Proc. of IEEE Global Telecommun. Conf.*, Miami, FL, Dec. 2010, pp. 1–5.

[25] A. Chaaban and A. Sezgin, "From non-causality to causality: degrees of freedom of the interference relay channel," in *Proc. of Conf. Info. Sciences Systems*, Princeton, NJ, Mar. 2010, pp. 1–6.

[26] T. Gou, S. A. Jafar, C. Wang, S.-W. Jeon, and S.-Y. Chung, "Aligned interference neutralization and degrees of freedom of the 2x2x2 interference channel," *IEEE Trans. Info. Theory*, vol. 58, no. 7, pp. 4381–4395, Jul. 2012.

[27] N. Lee and S. A. Jafar, "Aligned interference neutralization and the degrees of freedom of the 2 user interference channel with instantaneous relay," *submitted to IEEE Trans. Info. Theory*, vol. available at http://arxiv.org/abs/1102.3833, 2011.

[28] B. K. Chalise and L. Vandendorpe, "Optimization of MIMO relays for multipoint-to-multipoint communications: Nonrobust and robust designs," *IEEE Trans. Signal Processing*, vol. 58, no. 12, pp. 6355–6368, Dec. 2010.

[29] H. Ning, C. Ling, and K. K. Leung, "Relay-aided interference alignment: Feasibility conditions and algorithm," in *Proc. of IEEE Int. Symp. Info. Theory*, Austin, TX, Jun. 2010, pp. 390–394.

[30] L. Liang and G. Feng, "A game-theoretic framework for interference coordination in OFDMA relay networks," *IEEE Trans. Veh. Tech.*, vol. 61, no. 1, pp. 321–332, Jan. 2012.

[31] A. Zappone and E. A. Jorswieck, "Resource allocation in amplify-and-forward relay-assisted DS/CDMA systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 4, pp. 1271–1276, April 2012.

[32] Y. Shi, X. Dong, K. B. Letaief, and R. K. Mallik, "Coalition-assisted resource allocation in large amplify-and-forward cooperative networks," *IEEE Transactions on Vehicular Technology*, vol. 61, no. 4, pp. 1863–1873, May 2012.

[33] D. A. Schmidt, C. Shi, R. A. Berry, M. L. Honig, and W. Utschick, "Minimum mean squared error interference alignment," in *Proc. of Asilomar Conf. Signals, Systems Computers*, Pacific Grove, CA, Nov. 2009, pp. 1106–1110.

[34] S. W. Peters and R. W. Heath, Jr., "Interference alignment via alternating minimization," in *Proc. of IEEE Int. Conf. Acoustics, Speech Signal Processing*, Taipei, Taiwan, Apr. 2009, pp. 2445–2448.

[35] ——, "Cooperative algorithms for MIMO interference channels," *IEEE Trans. Veh. Tech.*, vol. 60, no. 1, pp. 206–218, Jan. 2011.

[36] D. P. Bertsekas, *Nonlinear programming*, 2nd ed. Athena Scientific, 1999.

[37] J. C. Bezdek and R. J. Hathaway, "Convergence of alternating optimization," *Journal on Neural, Parallel, and Scientific Computations*, vol. 11, no. 4, pp. 351–368, Dec. 2003.

[38] K. T. Truong and R. W. Heath, Jr., "Interference alignment for multiple-antenna amplify-and-forward relay interference channel," in *Proc. of IEEE Asilomar Conf. Signals, Systems Computers*, Pacific Grove, CA, USA, Nov. 2011, pp. 1–5.

[39] S. Ren and M. Van Der Schaar, "Distributed power allocation in multi-user multi-channel cellular relay networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 6, pp. 1952–1964, Jun. 2010.

[40] C. Raman, G. J. Foschini, R. A. Valenzuela, R. D. Yates, and N. B. Mandayam, "Half-duplex relaying in downlink cellular systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 5, pp. 1396–1404, May 2011.

[41] S. A. Jafar, K. S. Gomadam, and C. Huang, "Duality and rate optimization for multiple access and broadcast channels with amplify-and-forward relays," *IEEE Trans. Info. Theory*, vol. 53, no. 10, pp. 3350–3370, Oct. 2007.

[42] Y.-W. Liang and R. Schober, "Amplify-and-forward multi-antenna beamforming with joint source-relay power constraint," in *Proc. of IEEE Veh. Tech. Conf.*, Taipei, Taiwan, Sep. 2010, pp. 1–5.

[43] M. R. McKay, I. B. Collings, and A. M. Tulino, "Achievable sum rate of MIMO MMSE receivers: A general analytic framework," *IEEE Trans. Info. Theory*, vol. 56, no. 1, pp. 396–410, Jan. 2010.

[44] D. Guo, S. Shamai, and S. Verdu, "Mutual information and minimum mean-square error in Gaussian channels," *IEEE Trans. Info. Theory*, vol. 51, no. 4, pp. 1261–1282, Apr. 2005.

[45] S. S. Christensen, R. Agarwal, E. Carvalho, and J. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4792–4799, Dec. 2008.

[46] M. Razaviyayn, H. Baligh, A. Callard, and Z.-Q. Luo, "Joint transceiver design and user grouping in a MIMO interfering broadcast channel," in *Proc. of IEEE Conf. Info. Sciences Systems*, Princeton, NJ, Mar. 2011.

[47] S. Xu and Y. Hua, "Source-relay optimization for a two-way MIMO relay system," in *Proc. of IEEE Int. Conf. Acoustics, Speech Signal Processing*, Dallas, TX, Mar. 2010, pp. 3038–3041.

[48] K.-J. Lee, H. Sung, E. Park, and I. Lee, "Joint optimization for one and two-way MIMO AF multiple-relay systems," *IEEE Trans. Wireless Commun.*, vol. 9, no. 12, pp. 3671–3681, Dec. 2010.

[49] Q. Shi, M. Razaviyayn, Z. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Signal Processing*, vol. 59, no. 9, pp. 4331–4340, Sep. 2011.

[50] H. Lutkepohl, *Handbook of Matrices*.   John Wiley and Sons, Ltd., 1996.

[51] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*.   Press Syndicate of the University of Cambridge, 1991.

[52] Y. Huang and D. P. Palomar, "Rank-constrained separable semidefinite programming with applications to optimal beamforming," *IEEE Trans. Signal Processing*, vol. 58, no. 2, pp. 664–678, Feb. 2010.

[53] M. Grant and S. Boyd, "CVX: MATLAB software for disciplined convex programming," http://cvxr.com/cvx, Apr. 2011.

[54] W. Ali, Y. Huang, and S. Zhang, "New results on Hermitian matrix rank-one decomposition," *Math. Program., Ser. A*, vol. 120, no. 2, 2009.

[55] Y. Huang and S. Zhang, "Complex matrix decomposition and quadratic programming," *Math. Oper. Res.*, vol. 32, no. 3, pp. 758–768, 2007.

[56] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Processing Mag.*, vol. 27, no. 3, pp. 20–34, May 2010.

[57] R. A. Horn and C. R. Johnson, *Matrix Analysis*.   Cambridge University Press, 1990.

[58] M. V. Solodov, "On the convergence of constrained parallel variable distribution algorithm," *SIAM J. Optimization*, vol. 8, no. 1, pp. 187–196, Feb. 1998.

[59] V. R. Cadambe, S. A. Jafar, and C. Wang, "Interference alignment with asymmetric complex signaling-settling the Host-Madsen-Nosratinia conjecture," *IEEE Trans. Info. Theory*, vol. 56, no. 9, pp. 4552–4565, Sep. 2010.

[60] Z. Fang, Y. Hua, and J. C. Koshy, "Joint source and relay optimization for a non-regenerative MIMO relay," in *Proc. of IEEE Sensor Array Multichannel Processing*, Waltham, MA, Jul. 2006, pp. 239–243.

[61] Y. Rong, X. Tang, and Y. Hua, "A unified framework for optimizing linear nonregenerative multicarrier MIMO relay communication systems," *IEEE Trans. Signal Processing*, vol. 57, no. 12, pp. 4837–4851, Dec. 2009.